REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. AGENCY USE ONLY *(Leave blank)*	2. REPORT DATE	3. REPORT TYPE AND DATES COVERED
	7.Jan.99	THESIS

4. TITLE AND SUBTITLE
INTERCOMPARISION OF ICING AVIATION IMPACT VARIABLE FORECASTS PRODUCED DURING REAL-TIME MESOSCALE NUMERICAL WEATHER PREDICTION.

5. FUNDING NUMBERS

6. AUTHOR(S)
CAPT STOCK CHRISTOPHER M

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
UNIVERSITY OF OKLAHOMA

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)
THE DEPARTMENT OF THE AIR FORCE
AFIT/CIA, BLDG 125
2950 P STREET
WPAFB OH 45433

10. SPONSORING/MONITORING AGENCY REPORT NUMBER
FY99-31

11. SUPPLEMENTARY NOTES

12a. DISTRIBUTION AVAILABILITY STATEMENT
Unlimited distribution
In Accordance With AFI 35-205/AFIT Sup 1

12b. DISTRIBUTION CODE

13. ABSTRACT *(Maximum 200 words)*

14. SUBJECT TERMS

15. NUMBER OF PAGES

16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT

UNIVERSITY OF OKLAHOMA

GRADUATE COLLEGE

INTERCOMPARISON OF ICING AVIATION IMPACT VARIABLE FORECASTS PRODUCED DURING REAL-TIME MESOSCALE NUMERICAL WEATHER PREDICTION

A THESIS

SUBMITTED TO THE GRADUATE FACULTY

in partial fulfillment of the requirements for the

degree of

MASTER OF SCIENCE IN METEOROLOGY

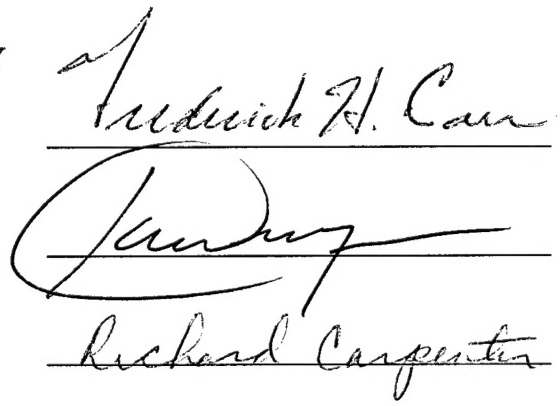
By

CHRISTOPHER MICHAEL STOCK
Norman, Oklahoma
1998

INTERCOMPARISON OF ICING AVIATION IMPACT VARIABLE FORECASTS PRODUCED DURING REAL-TIME MESOSCALE NUMERICAL WEATHER PREDICTION

A THESIS APPROVED FOR THE
SCHOOL OF METEOROLOGY

BY



ACKNOWLEDGEMENTS

The completion of my Master's Thesis is one of the highlights of my academic career. I can't begin to count the hours spent, and headaches suffered, getting through classes, writing computer code, reading endless papers and journal articles, and finally sitting down to write this thesis. I do know that I could never have done it alone. There is no way I can mention everyone who has had a hand in keeping me on track, but there are a few who I do need to single out.

I wish to pay special thanks my advisor, Dr. Fred Carr, for beginning the project that brought me back to OU. My interest in finding the best way to provide weather information to Air Force forecasters fit perfectly with his goals for the COMET-Tinker project. His continual support and guidance made it easy for me, an engineer by trade, to survive in the graduate weather world at OU. I also wish to thank the members of my thesis committee. Dr. Richard Carpenter provided me with invaluable help in the initial setup of our forecast production, and graciously handled every question I threw his way. Secondly, Dr. Kelvin Droegemeier has been a constant supporter of the verification efforts of our project.

The faculty and staff at the School of Meteorology is one of the finest group of people I have ever had the pleasure to work with. I would like to thank all the professors for their support and assistance over the last two years, Tom Condo for his help with all my computer questions, and especially Celia, Marcia, and Nancy for letting me vent whenever I reached a breaking point.

The friendships I have made while in Oklahoma will be a cherished part of me forever. To Chris, Jon, Dan, Brian, TJ, Jace, Abby, Tina, Brady, Britny, Rodger, Doug,

and Jim, thanks for always being there. The one person who I could not have done without is Eric Kemp. The time we spent hashing out forecast products, going over code, and tweaking *everything*, is more than anyone should ever have to endure around me. You're the best!

I would also like to thank Dan Weber, Doug Kennedy, Howard Johnson, Greg Thompson, and Gene Bassett for their help in obtaining additional observational data and helping with computer programming issues that arose.

Of course, the greatest moral support came from my family. The love and understanding I get on a daily basis is the most special thing in my life. Without you, none of this would have even mattered. I love you all! This thesis is dedicated to the memory of my grandfather, William Lenartowicz. Grandpa, I miss you.

Table of Contents:

List of Tables:

List of Illustrations:

Figure 3.6: Number of Pilot Reports (PIREPs) of icing during the WOP97 grouped by date for (a) all "Yes"/"No" icing reports and (b) "Yes" only reports (intensity indicated by color shading)

Figure 3.7: Sample AFWA RAOB Icing Algorithm forecast product. Intensity indicated by the color shading (1 – Trace, 2 – Light, 3 – Moderate), type indicated by text (c – Clear, r – Rime, and m – Mixed)

Figure 3.8: Illustration (partial Skew-T diagram) of the NCAR/RAP Icing Algorithm – (a) stratiform (freezing drizzle) icing scenario, (b) freezing rain icing scenario, and (c) stable and unstable icing scenarios. The solid lines are the vertical temperature (T) and dew point (T_d) profiles as indicated. (after Thompson et al. 1996)

Figure 3.9: Sample NCAR/RAP Icing Algorithm forecast product. Existence of icing indicated by the color shading, type indicated by text (W – Warm Stratus, F – Freezing Rain, S – Stable, and U – Unstable)

Figure 3.10: Sample LAPS Icing Algorithm forecast product. Existence of icing indicated by the color shading (1 – Light, 2 – Moderate, 3 – Heavy), frequency indicated by text (c – Continuous and i – Intermittent).

Figure 3.11: Threat (# of WORST PIREPs per unit area; x 10^6) (km^{-2}) calculated for each precipitation type and cloud cover amount category. Category abbreviations follow standard contractions (e.g. ZL is freezing rain, XOB is obscured sky cover, etc) (after Bernstein 1996)

Figure 3.12: Sample Stovepipe Icing Algorithm forecast product. Existence of icing indicated by the color shading, type indicated by text (G – General Icing, S – Supercooled Liquid Droplets).

Figure 3.13: Interactions of significant microphysical processes within supercooled clouds. (a) Relative rate of the production of meteors, normalized by wG (b) simplified model interactions. Number refers to (a) (Abbreviations listed in Table 3.8) (after Tremblay et al. 1996)

Figure 3.14: Sample Tremblay SLW Icing Algorithm forecast product. Existence of icing indicated by any color shading (wG – SDEP > 0).

Figure 3.15: Cloud microphysical processes considered in the ice microphysics parameterization scheme (after Xue et al. 1995; Lin et al. 1983)

Figure 3.16: Sample ARPS SLW Icing Algorithm forecast product. Existence of icing indicated by color shading ($q_c + q_r$ mixing ratio > 0.01 g kg^{-1} with temperature < 0 ° C).

Figure 4.41: Data in Table 4.3 in graphical format. Includes (a) total percent within accuracy criteria (%), (b) % total within accuracy criteria with a positive bias (%+), and (c) the % total within accuracy criteria with a negative bias (%-).

Figure 4.42: Average bias (left) and RMSE (right) for surface temperature (°C) for (a) 02, (b) 05, and (c) 08 hour forecast point (15, 18, 21 UTC respectively). (Negative values are indicated by dashed isopleths, while solid ones indicate positive values)

Figure 4.43: Same as Figure 4.42 except for surface dewpoint (°C).

Figure 4.44: Same as Figure 4.42 except for 850mb temperature (°C).

Figure 4.45: Same as Figure 4.42 except for 850mb dew point (°C).

Figure 4.46: Meteogram comparison for forecast on 24 December 1997 valid at (a) Lubbock, TX, and (b) Waco, TX.

Figure 4.47: Mean (a) POD_{yes} and (b) POD_{no} for all PIREPs over the entire WOP97.

Figure 4.48: Mean (a) Impacted Area (km^2) and (b) Impacted Volume (km^3) for entire forecast period.

Figure 4.49: Mean Area Efficiency (AE; $x10^{-6}$ km^{-2}) and Volume Efficiency (VE; $x10^{-6}$ km^{-2}) for entire forecast period.

Figure 4.50: Tremblay icing forecast at (a) a north-south cross-sections through the center of the domain and (b) an altitude of 18,000 feet (FL180) on 23 December 1997.

Figure 4.51: Average POD_{yes} for all icing forecast algorithms at each forecast hour.

Figure 4.52: Average Impacted Area (IA) for all icing forecast algorithms at each forecast hour.

Figure 4.53: Average Area Efficiency (AE) for all icing forecast algorithms at each forecast hour.

Figure 4.54: Average Volume Efficiency (VE) for all icing forecast algorithms at each forecast hour.

Figure 4.55: Daily (a) Impacted Area (IA) and (b) Area Efficiency (AE) encompassing all icing PIREPs

Figure 4.56: Daily (a) Impacted Volume (IV), and (b) Volume Efficiency (VE) encompassing all icing PIREPs

ABSTRACT

In 1996 a three year joint effort, Project COMET-Tinker, was initiated between the University of Oklahoma (OU), the Air Force Weather Agency (AFWA), and the Cooperative Program for Operational Meteorology, Education, and Training (COMET) to evaluate the use of real-time mesoscale numerical weather prediction (NWP) by USAF forecasting personnel. Its goal is to examine forecasts of specific aviation impact variables (AIVs) which could be incorporated directly into the weather services provided by the base weather station (BWS) at Tinker Air Force Base (AFB).

During a Winter Operational Period (WOP), 23 December 1997 to 31 January 1998, daily nine hour, nine kilometer resolution forecasts were made using the Advanced Regional Prediction System (ARPS) developed by the Center for Analysis and Prediction of Storms (CAPS). Icing forecast products derived from algorithms developed by various weather agencies world-wide were generated using the model output and disseminated to the Tinker BWS via the world wide web (WWW).

Verification procedures developed at OU utilizing observational data from surface reporting stations, the Oklahoma Mesonet, and wind profilers, along with comparisons against upper air analysis fields from the Rapid Update Cycle (RUC) were used to evaluate general model performance. Forecast errors were found to be consistent with known model deficiencies and limited to the lower portions of the model domain. Pilot Reports (PIREPs) of in-flight icing conditions were then compared with the ARPS derived icing forecasts to determine the effectiveness of ARPS and the algorithms utilized in predicting the observed icing conditions. Evaluation of the icing forecasts

showed that "synoptically" based algorithms performed better than those that relied on microphysical parameterizations inherent in ARPS and other mesoscale models.

Chapter 1: Introduction

Improvements in the field of numerical weather prediction have led to an increase in the overall utility of various weather modeling systems. Research efforts have focused on the development of accurate numerical techniques within the models themselves, the inclusion of explicit treatments of various atmospheric processes that take place on sub-grid scales, the efficient incorporation of all available observational data sources into the model initialization, and the evaluation of model performance against observed weather conditions. One aspect of mesoscale model development that has received a good deal of interest recently is in the ability to provide explicit forecast information on weather variables needed by operational weather forecasters.

Typically the predicted values of individual meteorological variables act as the basis for the display of numerical weather prediction results to operational forecasters. The ability to provide forecast information for specific weather phenomena of interest would greatly increase the usefulness of the model forecasts. The development of specific forecast products for the commercial and military aviation communities, called aviation impact variables (AIVs), are based on the mission requirements of the end user of the forecast data. The needs within the United States Air Force (USAF) weather community are for AIVs related to the in-flight operation of aircraft and for the protection of critical resources on the ground. Since many of these weather elements are not directly forecast by the model, procedures must be developed to relate the forecast output to the desired weather element. One such AIV desired by the USAF is the indication of in-flight icing conditions. Aircraft icing has been shown to have significant impact on

aircraft operations. Therefore the ability of a forecast model to provide direct indications of icing conditions within the forecast domain is of interest.

It is expected that with the advances made in mesoscale modeling, and the continuing research being conducted on the atmospheric conditions associated with aircraft icing, that current mesoscale models should have some success in forecasting icing conditions. The focus of this work is to test the ability of the Advanced Regional Prediction System (ARPS) mesoscale model, developed by the Center for Analysis and Prediction of Storms (CAPS), to produce accurate icing aviation impact variable forecasts for use in the real time operations of the base weather station at Tinker Air Force Base, Oklahoma. Daily forecasts using ARPS were made during the winter of 1997-98. During this period, specific indications of forecast icing conditions were made using several different computational processes. The ability of these algorithms to correctly predict regions where icing was observed was compared using a combination of traditional verification measures and those related specifically to the use of pilot reports (PIREPs) in the verification of observed weather phenomena.

In addition to the verification of icing forecasts produced, the overall ability of the ARPS model to predict the future state of the atmosphere was examined to ensure the icing algorithms had accurate forecast information on which to perform their calculations. The quality of the model output was compared using various sources of surface and upper-air observations of temperature, dew point, wind speed and direction, and other standard meteorological variables.

Chapter 2 Background on Project COMET-Tinker

Section 2.1 Overview of Tinker Air Force Base Weather Station Operations

The first step in understanding the rationale behind the forecasting efforts utilized for this research is to obtain a working knowledge of the operational needs within a typical Air Force Base Weather Station (BWS). The specific aircraft assigned to a base, their specialized support organization, the wide range of unique missions, and the underlying base support structure, all establish a unique set of weather related requirements that must be met by the local forecasters. There however, exists a core set of weather information responsibilities, which must be fulfilled by the station regardless of the specialized units assigned. In addition, we must also be cognizant of the resources, both personnel and equipment, that are at the disposal of the BWS to accomplish its mission. While the forecasting efforts made by the COMET-Tinker project were initially established for the BWS at Tinker Air Force Base (Tinker AFB), Oklahoma, these activities can be easily applied to the needs of weather stations around the Air Force. The following discussion of the operation and requirements of the BWS at Tinker will be used as our guide to identify the needs this project was designed to fill.

2.1.1 Personnel

At Tinker Air Force Base, the 72nd Operations Support Squadron, Base Weather Station, (also referred to as the "weather station" or just "station" for remainder of the discussion) is the primary agency responsible for addressing the weather needs of the Tinker community. The weather station consists of 18 active duty Air Force members, and one civilian forecaster.

The station leadership is headed by the Flight Commander (command meteorologist), typically a Captain with at least seven to nine years experience in the weather career field. Overseeing the entire operations of the weather station, the commander is the primary person responsible for ensuring that the weather service needs of the base are met. The commander is the main weather point of contact for all supported units and the base populace. The Wing Weather Officers (operational meteorologists), normally in the ranks of Lieutenant (Second and/or First) with typically less than four years forecasting experience, are focal point for the flying units assigned to Tinker. They deal with the day-to-day flying mission requirements relating to the weather station. They conduct regular safety training on weather hazards experienced by pilots and ground crew, provide routine weather briefing support to command and control elements (staff support), arrange specialized weather support requirements for exercises and deployments, coordinate other routine weather support needs, and assist the commander in integrating all identified requirements into the daily operations of the station. Finally, the Superintendent of Weather Operations (SWO), the senior enlisted

member of the station (typically a Master Sergeant (MSgt)), is responsible for the final integration and actual implementation of all procedures dealing with the weather observing and forecasting functions of the station. The SWO, usually someone with six to eight years of forecasting and BWS experience, provides the backbone of forecasting knowledge for the younger and less experienced duty forecasters. In addition, the SWO works with the commander and WWOs to ensure that the products generated and procedures in place within the weather station meet the regular needs of their customers.

The remainder of the station personnel consists of the duty forecasters and observers, operating on rotating shifts of eight hours each. These station members produce the majority of daily weather information products used by the weather station customers. They interact on a continual basis with all base agencies that require weather support.

The Tinker BWS has five duty observers of the rank of either Airman First Class (A1C) or Senior Airman (SrA). They are the entry level enlisted ranks within the Air Force and have likely just completed their initial basic training and indoctrination in the Air Force. Upon completing basic training they are sent to a three months initial observing skills training course at Keesler AFB, MS, before being assigned to their first duty station. Observers at Tinker complete in-house training on station specific mission requirements, unique weather features of the region, the observing equipment available, and other procedural elements of station operation before becoming certified observers. Certification of all observers is renewed annually to ensure the most accurate and timely observations possible.

The station has seven active duty forecasters, ranging in rank from SrA to Technical Sergeant (TSgt), and one civilian forecaster assigned. The active duty personnel have anywhere from six months to over six years experience. Most Air Force forecasters complete three to four years as an observer before continuing on into the forecasting career path. New forecasters are also initially trained in an intensive six to nine month period at Keesler AFB. At Keesler they receive comprehensive instruction on atmospheric dynamics, synoptic meteorology, and general forecasting tools and techniques. The training is an overview of what an undergraduate sees in their four years of university education with emphasis on Air Force needs and procedures. Once assigned to Tinker, the forecaster also completes in-house training, including additional orientation on specific weather regimes and forecast considerations of the Central and Southern Plains, before becoming a certified duty forecaster. All forecasters at Tinker also maintain dual certification as an observer.

Finally, the civilian forecaster is assigned to provide continuity for the ever-changing mix of forecasters assigned to the station. As new forecasters replace those leaving, the valuable experience gained specific to Tinker is maintained by the civilian forecaster who can pass the lessons learned from the past weather situations to the new forecasting personnel.

2.1.2 Equipment

To accomplish their mission, weather station personnel have a vast array of meteorological and computer equipment at their disposal. For continuous and automated

observational elements the standard BWS equipment includes sensors to measure the local temperature, dew point, wind speed and direction, pressure, ceiling, and runway visual range. The actual components used to make these continuous measurements are listed in Table 2.1. The locations of these sensors are displayed in Figure 2.1, as well as the "Official" observing point where the manual assessment of cloud cover, visibility and related obstructions, precipitation type and amounts, and other observational elements are determined. The observing personnel also have various manual observing equipment such as sling-psychrometers to use in the event of automated equipment outages. Finally many equipment items considered as "forecasting" tools such as the lightning detection display and the weather radar are used as sources for observational elements. Most notably they assist in the determination of thunderstorm location and movement, the presence of low-level wind shear, and other items routinely encoded in the remarks section of surface observations.

On the forecasting side of the station, there is an overwhelming amount of weather data that are available to the duty forecaster. The fundamental source of weather information at the BWS is the Automated Weather Distribution System (AWDS). AWDS acts as a the primary data feed between the Air Force Weather Agency (AFWA), located at Offutt AFB, NE, and all weather stations around the world. Through the AWDS system the duty forecaster is able to access raw surface and rawinsonde observations from around the world, forecast products produced in-house at Offutt AFB by Air Force personnel looking at the synoptic scale features, model and forecast products produced by the National Weather Service, along with a great deal more. Through AWDS, it is possible to display and analyze surface and standard pressure level

Name	Sensor	Function
Laser Beam Ceilometer	GMQ-34	Measures cloud height (hundred of feet) when they exist below 12,000 feet above ground level (AGL)
Temperature and Dewpoint Set	FMQ-8	Measures the local ambient temperature and dew point (degrees Celsius)
Digital Wind Sensor	FMQ-13	Measures a 2 minute running average (updated every 5 seconds) and peak (past 10minute/60minute/24hour) values of the wind speed and direction to the nearest whole knot (Four sensors, one located at both ends of each runway.)
Digital Barometer and Altimeter Setting Indicator (DBASI)	ML-658	Measures the local station pressure (millibars) and altimeter setting (inches of mercury)
Transmissometer	GMQ-32	Measures the Runway Visual Range (RVR – feet)

Table 2.1: Description of weather station observing equipment.

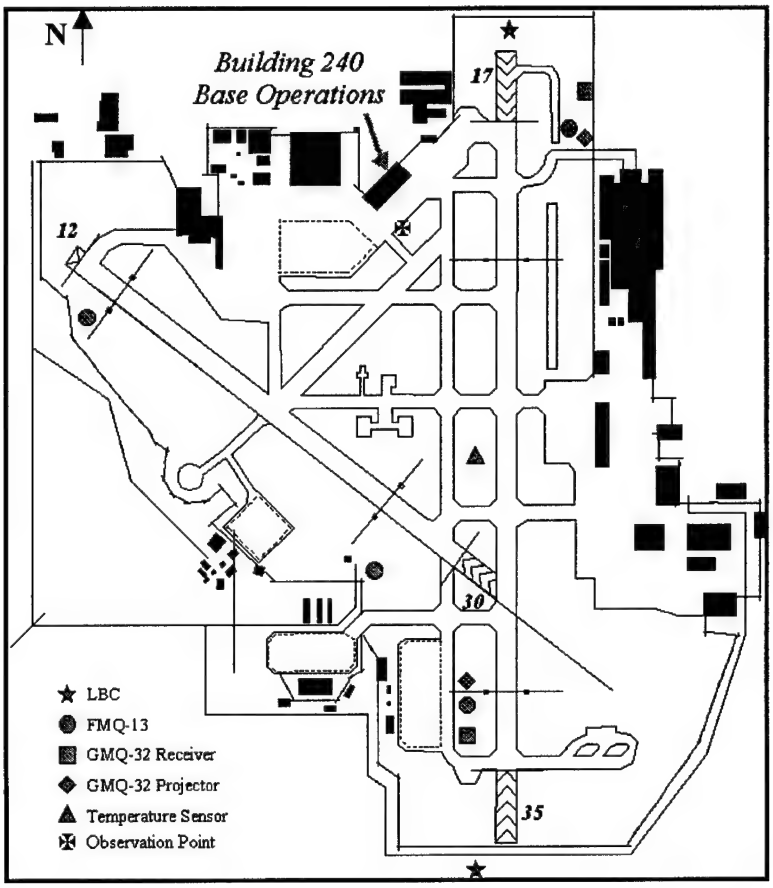


Figure 2.1: Location of weather sensors on Tinker Air Force Base. The location of the Base Weather Station (BWS), situated in Building 240, is also indicated as well as the "official" observing point, ✠.

observational and model data. AWDS also acts as the primary dissemination system for Tinker's observational and forecast products to the base agencies, as well as back to AFWA for availability to all other Air Force weather stations units.

Additional sources of forecast and observational information used in preparation of forecasting products include: the Weather Service Radar WSR-88D; a satellite image retrieval and display system from Harris Inc.; an Alden national lightning display system with data received from the GeoMet Data Services' (GDS) National Lightning Detection Network; access to data from the Oklahoma Mesonet operated by the Oklahoma Climatological Survey (OCS); AFWA products through dial-up access to the Air Force Dial-In System (AFDIS) and over the World Wide Web (WWW) via the Air Force Weather Information Network (AFWIN); PCGridds data display capabilities; data from the Naval Oceanographic Data Distribution System (NODDS), and the large volume of additional WWW servers providing forecast and observational data in real time.

2.1.3 Units/Missions Supported

As indicated earlier, the mission requirements placed on a BWS are highly dependent on the types of units that operate from the particular Air Force Base. The diverse community which occupies Tinker ranges from operational flying organizations and their related support units, to deployment-ready support forces, to military and contract maintenance functions. The following is a brief description of the major units supported by the Tinker AFB weather station, and their unique operational activities.

The main unit assigned to Tinker AFB is the Oklahoma City Air Logistics Center (OC-ALC) operated by the Air Force Material Command (AFMC), headquartered out of Wright-Patterson AFB, Ohio. AFMC is responsible for procurement and maintenance of Air Force aircraft and other combat systems. The logistics center is one of five depot repair facilities run by the Air Force and is the worldwide manager for a wide range of aircraft, engines, missiles, and commodity items. The center manages an inventory of over twenty-two hundred aircraft which include the B-1, B-2, B-52, C/KC-135, E-3, VC-25, VC-137, and other aircraft. Aircraft are brought from operating locations around the world to Tinker for comprehensive engine and structural repairs, repainting, routine preventative maintenance and refurbishing, system upgrade and augmentation, corrosion prevention, and other command level maintenance which cannot be completed by each individual unit.

Each aircraft must be completely evaluated and flight-tested before being returned to active service. The primary flying component of the center, the 10th Test Flight, is responsible for ensuring that all aircraft leaving the center satisfactorily complete a thorough set of ground and in-flight operation checks. Specific "no-fly" weather conditions must be adhered to depending on the aircraft being evaluated. The spatial forecast area requirements vary for each aircraft type. They generally involve flying operations covering a significant portion of the Southern Plains region.

The OC-ALC is also responsible for the day to day operations of Tinker AFB. The 72nd Air Base Wing is the host organization for Tinker responsible for all base and infrastructure support. The base is made up of nearly eight hundred buildings covering over five thousand acres of land. The two active runways are surrounded by over ninety

acres of ramp and hangar space. The Wing supports the entire installation, providing critical functions such as security, fire protection, medical services, civil engineering, communications, supply, and airfield operations. The base is comparable to a city with a population of over 30,000.

There are two active duty flying units assigned to Tinker AFB. The first is the 552nd Air Control Wing (ACW) which operates the E-3 airborne warning and control system (AWACS) aircraft, shown in Figure 2.2. The ACW is part of the Air Force's Air Combat Command (ACC), headquartered out of Langley AFB, Virginia, responsible for worldwide operations of combat and support aircraft for the Air Force. The E-3 is a modified Boeing 707 commercial airframe over 145 feet (44 meters) long, 40 feet (12 meters) tall, with a wingspan covering 135 (40 meters). A sophisticated radar, over thirty

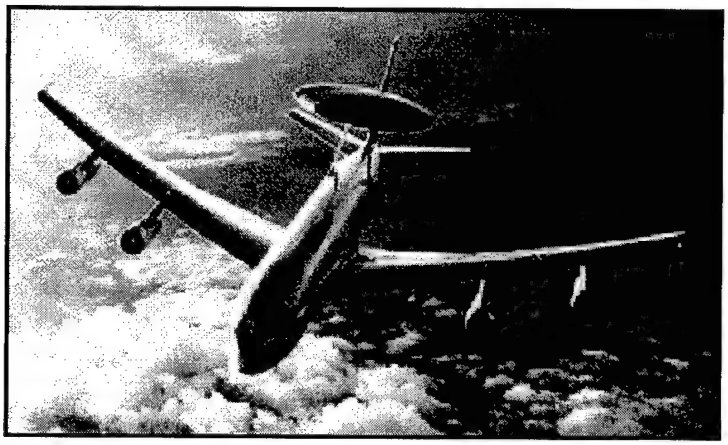


Figure 2.2: E-3 airborne warning and control system (AWACS) aircraft.

feet in diameter, six feet thick, and supported by struts eleven feet above the airframe, is mounted on its back. The radar and other on-board sensors provide all weather surveillance, warning, interception control, and airborne battle management capabilities for Department of Defense (DoD) activities. During conflicts the AWACS acts as a

mobile airborne platform monitoring enemy and friendly aircraft, naval vessels, and ground forces. It provides direct information to combat aircraft for reconnaissance, air support of ground activities, and air interdiction operations. In addition it provides critical information to rear area command and control locations.

Typical AWACS flights consist of flying a particular location ("station") for extended periods. These stations can comprise regions of thousands of square miles. Flight altitudes are usually above twenty five thousand feet, with missions lasting upwards of twelve hours. While "on station" the aircraft circles ("orbits") the specified location while on-board sensors are monitored for mission specified activities within a designated area of interest (control areas). Within these "control areas", the AWACS crew is responsible for coordinating the activities of all aircraft in the assigned area.

Of special interest to the Tinker weather station are the regular training missions that actually depart from the base. These training missions are conducted at standard orbit/control locations throughout the continental United States (CONUS). These are illustrated in Figure 2.3. For missions departing Tinker, the majority stay within the

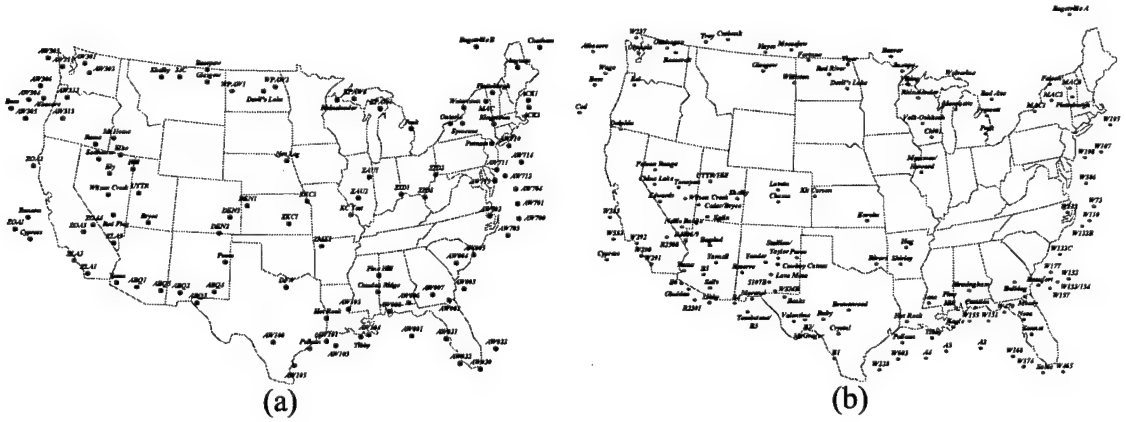


Figure 2.3: 552nd Air Control Wing routine Continental United States (CONUS) training (a) "Orbit" and (b) "Control" areas.

Central and Southern Plains, but can reach into the far Southwest and Southeast sections of the US. Due to the precarious nature of the radar dome, the E-3 has extensive weather needs, especially related to protecting the aircraft from significantly high winds on the ground, hail and severe thunderstorm activity, and in flight icing and turbulence.

The other active duty flying unit assigned to Tinker is the United States Navy's (USN) Strategic Communications Wing One (STRATCOMWINGONE) which operates two Fleet Air Reconnaissance Squadrons (VQ3 and VQ4). The Navy wing at Tinker also flies a derivative of the Boeing 707, the E-6 Mercury, shown in Figure 2.4. There are sixteen E-6 aircraft assigned to STRATCOMWINGONE.

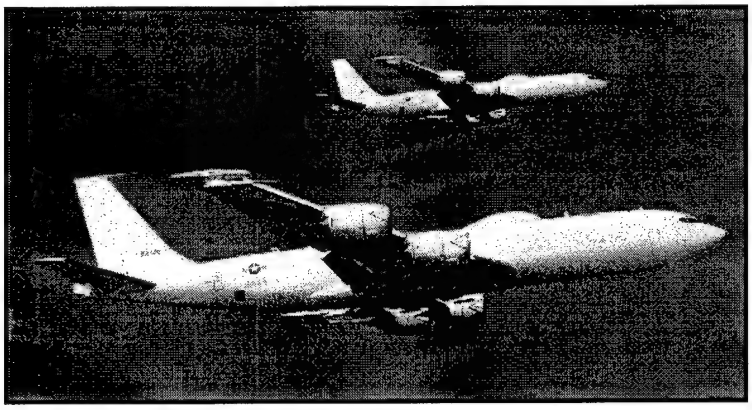


Figure 2.4: E-6 communications platform aircraft.

A one-of-a-kind unit in the DoD, the E-6 was originally designed to ensure reliable communications capabilities with the submerged ballistic missile submarine fleet in the event of a war. The wing's responsibilities have broadened to encompass providing a secure communications link between all naval vessels around the world and centralized command and control agencies. Using antennas trailed from the rear of the aircraft, the E-6 relays information received from various ground and satellite-based

sources to deployed naval vessels. A return communications link back to the command and control elements is also maintained. The wing is also in the process of taking over the Airborne Command Post (ABNCP) and National Command Authority (NCA) communication and operations responsibilities currently handled by the EC-135 Looking Glass aircraft. The E-6 is considered to be the primary source of electro-magnetic pulse hardened, mobile, command and control capability for national defense. Similar to the E-3 AWACS, the E-6 flies out to a specific location and "orbits" for an extended time period. This enables the continuous communication with deployed naval forces and/or capabilities for a non-vulnerable command and control platform. Flights originating from Tinker operate throughout the CONUS.

Tinker is also home to Oklahoma's only Air Force Reserve (AFR) flying unit, the 507th Air Refueling Wing (ARW). The wing flies the KC-135 Stratotanker aircraft shown in Figure 2.5, which is also manufactured by the Boeing Company. The KC-135

Figure 2.5: KC-135 air-refueling aircraft shown with refueling "boom" extended.

is similar in dimensions to the E-3 and E-6 aircraft. The wing supports U.S. military and North Atlantic Treaty Organization (NATO) aircraft operations with aerial refueling

missions world-wide. While they are worldwide capable, the 507th routinely flies its operational and training missions of interest to the Tinker BWS only within the CONUS. The refueling missions are also similar to AWAC and STRATCOMWINGONE activities in that the mission profile requires them to orbit a specified locale where they meet up with additional aircraft. Once contact is made they provide fuel through a "boom" extending from the rear of the aircraft, enabling the other aircraft to continue operations without landing. The normal "on station" time is less than that for the two active duty units at Tinker, but one KC-135 aircraft may make multiple refueling contacts at several different locations.

The final major unit assigned to Tinker is the 3rd Combat Communications Group, commonly referred to as the 3rd Herd. The 3rd Herd is also an ACC operated unit that is responsible to provide deployable communications, computer systems, navigational aids, and air traffic control services anywhere in the world in support of Air Force, DoD, and other U.S. commitments. The over eight hundred personnel assigned are trained to deploy more than one-hundred fifty separate mission systems to provide initial service to locations where these capabilities do not exist.

The remaining base and Air Force level specialized units assigned to Tinker coordinate various essential peacetime and wartime support operations. These range from Air Force wide accounting and military payroll offices, centralized computer and communication facilities, deployment capabilities, and other support activities. Most rely on the weather station for routine weather observational, forecast, and severe weather notification information.

The last group of agencies that the weather station supports consists of flying units not located at Tinker. These include the Army Aviation Support Facility in Lexington, Oklahoma, which operates UH-1 and UH-60 helicopters; the 137th Airlift Wing in Oklahoma City, OK, which operates C-130 transport aircraft; the 188th Air National Guard Fighter Wing at Fort Smith, AR, which flies F-16 fighters; and a significant number of transient aircraft which temporarily operate out of the base. The transient aircraft range throughout the entire DoD inventory from small training aircraft, to fighter aircraft, up to large airframe transports and bombers, and can generate as many as two hundred additional flight weather briefings a month. While numerous, the support needed for these off-base customers is similar to various aspects support provided to assigned units, or to flight weather briefing support that is standard at all Air Force weather stations.

2.1.4 Weather Station Services

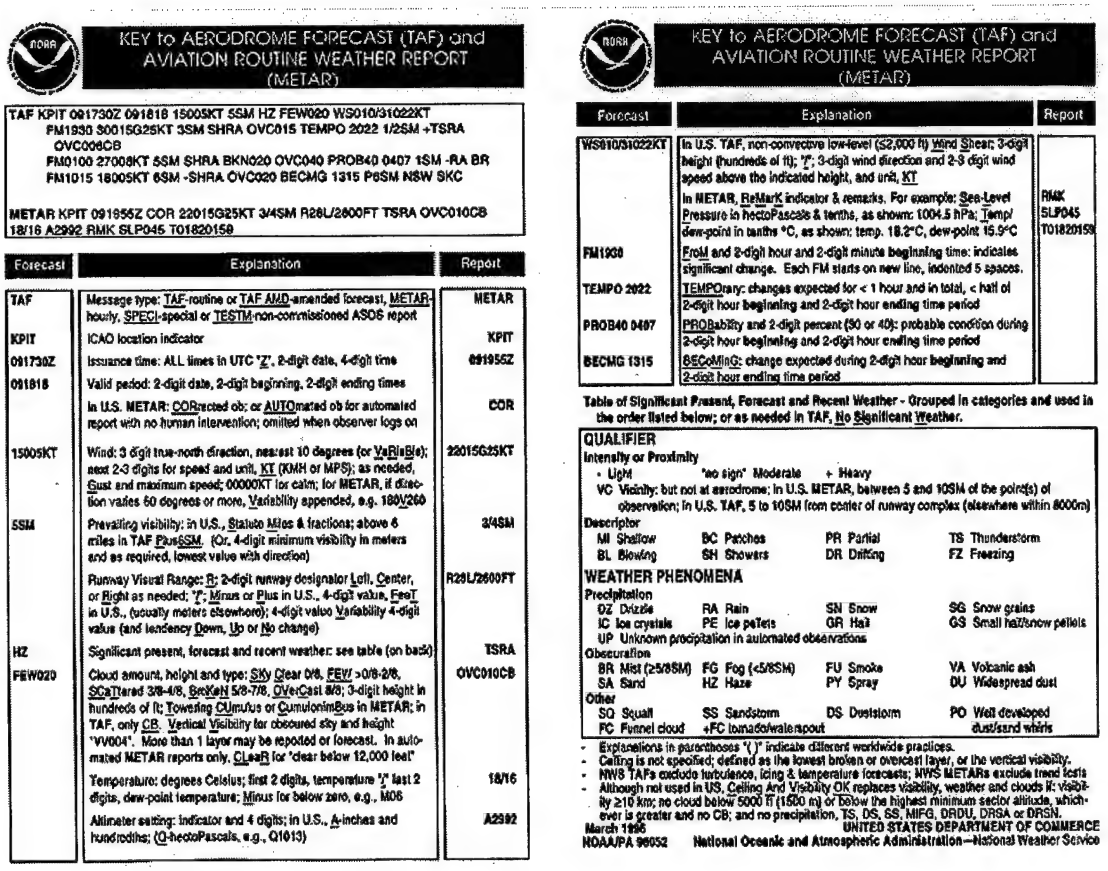
To provide the necessary weather information to fulfill the operational requirements of supported units, the weather station has a standard set of forecasting and observing products produced on a regular basis. Various DoD, Air Force, base level, and other local publications as outlined in the station's Weather Support Plan (Blaine 1996) dictate these services. The duty forecaster and observer, along with other station personnel work in tandem to ensure all products and information leaving the station are consistent, and reflect the best possible product (i.e. forecast) that can be achieved. Since the COMET-Tinker project is geared towards specific elements of the station's

forecasting activities, this section will highlight those that are directly impacted by the forecast data provided. Although essential to the complete operational effectiveness of the station, activities such as the taking of hourly surface observations and the preparation of staff support information will not be discussed here.

The main service provided by the weather station is the preparation and dissemination of a Terminal Aerodrome Forecast (TAF). Figure 2.6 shows a sample TAF. While Figure 2.7 provides a description of elements and breakdown of the decoding information. The TAF provides a twenty-four hour forecast of essential weather elements to all agencies, on and off base, and other activities needing detailed descriptions of conditions expected at Tinker for the specified time period. The TAF is designed to encompass, to the nearest hour, the changes of weather conditions across pre-determined thresholds. Conditions which are included are wind direction and speed, visibility, sky cover amounts and altitude, altimeter setting, turbulence and icing below ten thousand feet mean sea level (MSL), and various significant weather information such as the presence of non-convective low-level wind shear (LLWS), thunderstorms in the vicinity (within 25 nautical miles), or other features. The forecast is issued every six hours at 00, 06, 12 and 18 Universal Time (UTC) covering the next twenty-four hour period. If the forecast conditions change or are expected to change from those reflected in the current forecast, an amendment (AMD) must be issued to reflect the current or expected conditions. Table 2.2 lists the forecast and amendment criteria for Tinker AFB.

KTIK FCST 23-23 17012G22KT 7 FEW070 SCT150 SCT300 LGT TURB SFC-040 ALSTG29.94INS
BECMG 04-05 16009KT 7 SCT060 BKN120 BKN300 ALSTG29.98INS
BECMG 14-15 18012G20KT 7 SCT045 SCT140 BKN300 LGT TURB SFC-040 ALSTG29.94INS
BECMG 21-22 20012G24KT 7 SCT045CB BKN120 BKN280 VCTS ALSTG29.90INS

Figure 2.6: Sample Terminal Aerodrome Forecast (TAF)



KEY to AERODROME FORECAST (TAF) and AVIATION ROUTINE WEATHER REPORT (METAR)

TAF KPIT 091730Z 091818 15005KT 5SM HZ FEW020 WS010/31022KT
FM1930 30015G25KT 3SM SHRA OVC015 TEMPO 2022 1/2SM +TSRA OVC006CB
FM0100 27008KT 5SM SHRA BKN020 OVC040 PROB40 0407 1SM -RA BR
FM1015 18005KT 6SM -SHRA OVC020 BECMG 1315 P6SM NSW SKC

METAR KPIT 091955Z COR 22015G25KT 3/4SM R28L/2600FT TSRA OVC010CB 18/16 A2992 RMK SLP045 T01820159

Forecast	Explanation	Report
TAF	Message type: TAF-routine or TAF AMD-amended forecast, METAR-hourly, SPECI-special or TESTM-non-commissioned ASOS report	METAR
KPIT	ICAO location indicator	KPIT
091730Z	Issuance time: ALL times in UTC "Z", 2-digit date, 4-digit time	091955Z
091818	Valid period: 2-digit date, 2-digit beginning, 2-digit ending times	
	In U.S. METAR: CORrected ob; or AUTOmated ob for automated report with no human intervention; omitted when observer logs on	COR
15005KT	Wind: 3 digit true-north direction, nearest 10 degrees (or VaRiaBle); next 2-3 digits for speed and unit, KT (KMH or MPS); as needed, Gust and maximum speed; 00000KT for calm; for METAR, if direction varies 60 degrees or more, Variability appended, e.g. 180V260	22015G25KT
5SM	Prevailing visibility: in U.S., Statute Miles & fractions; above 6 miles in TAF P6SM. (Or, 4-digit minimum visibility in meters and as required, lowest value with direction)	3/4SM
	Runway Visual Range: R; 2-digit runway designator Left, Center, or Right as needed; "/"; Minus or Plus in U.S., 4-digit value, FeeT in U.S., (usually meters elsewhere); 4-digit value Variability 4-digit value (and tendency Down, Up or No change)	R28L/2600FT
HZ	Significant present, forecast and recent weather: see table (on back)	TSRA
FEW020	Cloud amount, height and type: SKy Clear 0/8, FEW >0/8-2/8, SCaTtered 3/8-4/8, BroKeN 5/8-7/8, OVerCast 8/8; 3-digit height in hundreds of ft; Towering CUmulus or CumulonimBus in METAR; in TAF, only CB. Vertical Visibility for obscured sky and height "VV004". More than 1 layer may be reported or forecast. In automated METAR reports only, CLeaR for "clear below 12,000 feet"	OVC010CB
	Temperature: degrees Celsius; first 2 digits, temperature "/" last 2 digits, dew-point temperature; Minus for below zero, e.g., M06	18/16
	Altimeter setting: indicator and 4 digits; in U.S., A-inches and hundredths; (Q-hectoPascals, e.g., Q1013)	A2992

KEY to AERODROME FORECAST (TAF) and AVIATION ROUTINE WEATHER REPORT (METAR)

Forecast	Explanation	Report
WS010/31022KT	In U.S. TAF, non-convective low-level (≤2,000 ft) Wind Shear; 3-digit height (hundreds of ft); "/"; 3-digit wind direction and 2-3 digit wind speed above the indicated height, and unit, KT	
	In METAR, ReMarK indicator & remarks. For example: Sea-Level Pressure in hectoPascals & tenths, as shown: 1004.5 hPa; Temp/dew-point in tenths °C, as shown: temp. 18.2°C, dew-point 15.9°C	RMK SLP045 T01820159
FM1930	FroM and 2-digit hour and 2-digit minute beginning time: indicates significant change. Each FM starts on new line, indented 5 spaces.	
TEMPO 2022	TEMPOrary: changes expected for < 1 hour and in total, < half of 2-digit hour beginning and 2-digit hour ending time period	
PROB40 0407	PROBability and 2-digit percent (30 or 40): probable condition during 2-digit hour beginning and 2-digit hour ending time period	
BECMG 1315	BECoMinG: change expected during 2-digit hour beginning and 2-digit hour ending time period	

Table of Significant Present, Forecast and Recent Weather - Grouped in categories and used in the order listed below; or as needed in TAF, No Significant Weather.

QUALIFIER

Intensity or Proximity
- Light "no sign" Moderate + Heavy
- VC Vicinity: but not at aerodrome; in U.S. METAR, between 5 and 10SM of the point(s) of observation; in U.S. TAF, 5 to 10SM from center of runway complex (elsewhere within 8000m)

Descriptor

MI Shallow	BC Patches	PR Partial	TS Thunderstorm
BL Blowing	SH Showers	DR Drifting	FZ Freezing

WEATHER PHENOMENA

Precipitation

DZ Drizzle	RA Rain	SN Snow	SG Snow grains
IC Ice crystals	PE Ice pellets	GR Hail	GS Small hail/snow pellets
UP Unknown precipitation in automated observations			

Obscuration

BR Mist (≥5/8SM)	FG Fog (<5/8SM)	FU Smoke	VA Volcanic ash
SA Sand	HZ Haze	PY Spray	DU Widespread dust

Other

SQ Squall	SS Sandstorm	DS Duststorm	PO Well developed dust/sand whirls
FC Funnel cloud	+FC tornado/waterspout		

- Explanations in parentheses "( )" indicate different worldwide practices.
- Ceiling is not specified; defined as the lowest broken or overcast layer, or the vertical visibility.
- NWS TAFs exclude turbulence, icing & temperature forecasts; NWS METARs exclude trend forecasts
- Although not used in US, Ceiling And Visibility OK replaces visibility, weather and clouds if: visibility ≥10 km; no cloud below 5000 ft (1500 m) or below the highest minimum sector altitude, whichever is greater and no CB; and no precipitation, TS, DS, SS, MIFG, DRDU, DRSA or DRSN.

March 1996 UNITED STATES DEPARTMENT OF COMMERCE
NOAA/PA 96052 National Oceanic and Atmospheric Administration—National Weather Service

Figure 2.7: Terminal Aerodrome Forecast (TAF) standard description and code breakdown.

Forecast Threshold Criteria
Ceiling(feet)/ Visibility (statute miles) Category *Ceiling* *Visibility* 3000 3 1500 2 1000 ½ 200
Wind Speed change of 10 knots or more
Wind Direction change of 30 degrees when speeds/gusts are above 15 knots
Any Precipitation
Any Thunderstorms
Warning or Advisory condition expected
Icing or Turbulence below 10,000 feet not associated with thunderstorms
Low-level Wind Shear not associated with thunderstorms

(a)

Amendment Criteria
Ceiling(feet)/ Visibility (statute miles) Category: *Ceiling* *Visibility* 3000 or 3 1000 or 2 200 or ½
Freezing precipitation starts/stops
Operationally significant occurrence or non-occurrence of precipitation
Precipitation starts/stops causes warning or advisory to be issued, amended, or canceled
Compatibility with Warning/Advisory
Wind Speed error (speed/gust) of 10 knots or more or Wind Direction error (speed/gust) of 30 degrees or more when winds are greater than 15 knots
Beginning or ending of moderate (or severe) Turbulence not specified correctly in forecast
Beginning or ending of light or greater Icing not specified correctly in forecast
Low-level Wind Shear starts/stops

(b)

Table 2.2: Tinker Air Force Base (a) Terminal Aerodrome Forecast (TAF) and (b) amendment criteria.

As noted earlier, for each forecast shift the duty forecaster has a great deal of weather information, both observational and forecast based, to use in preparation of the next TAF. To ensure consistent analysis and forecast methodology between different forecasters, and to provide a forecast production framework for inexperienced forecasters, a TAF Worksheet (Figure 2.8) was created. The worksheet gives specific things to look for and analyze given the data available. A worksheet must be completed for every forecast issued by the station. Once the forecast is issued, the forecaster maintains a continuous watch on those elements that may change unexpectedly and uses the worksheet to document amendments issued until the next regularly scheduled forecast goes out. Finally, the worksheet allows for the case study examination of unforecast or incorrectly forecast weather events after the fact.

The second service provided by the weather station is to brief flight and mission weather information to pilots and aircrew members. To the flying customers, the primary source of weather information is the Department of Defense (DD) Form 175-1, Flight Weather Briefing, illustrated in Figure 2.9. All aircraft crews must receive a flight weather briefing prior to departure. In a typical month the duty forecasters at Tinker will see upwards of five to seven hundred flight briefing requests. The form covers all aspects of a standard aircraft flight and is used by all aircraft activities within the DoD. The top section, Part I (items 1-13) includes information relative to the departure from Tinker AFB, or whatever location appropriate. Part II, items 14-24, deal with specific en-route hazards such as icing, thunderstorms, turbulence, and visibility restrictions, and must indicate regions of varying conditions or when conditions will change as the flight progresses. Part III list forecasts for the arrival time at locations desired by the pilot,

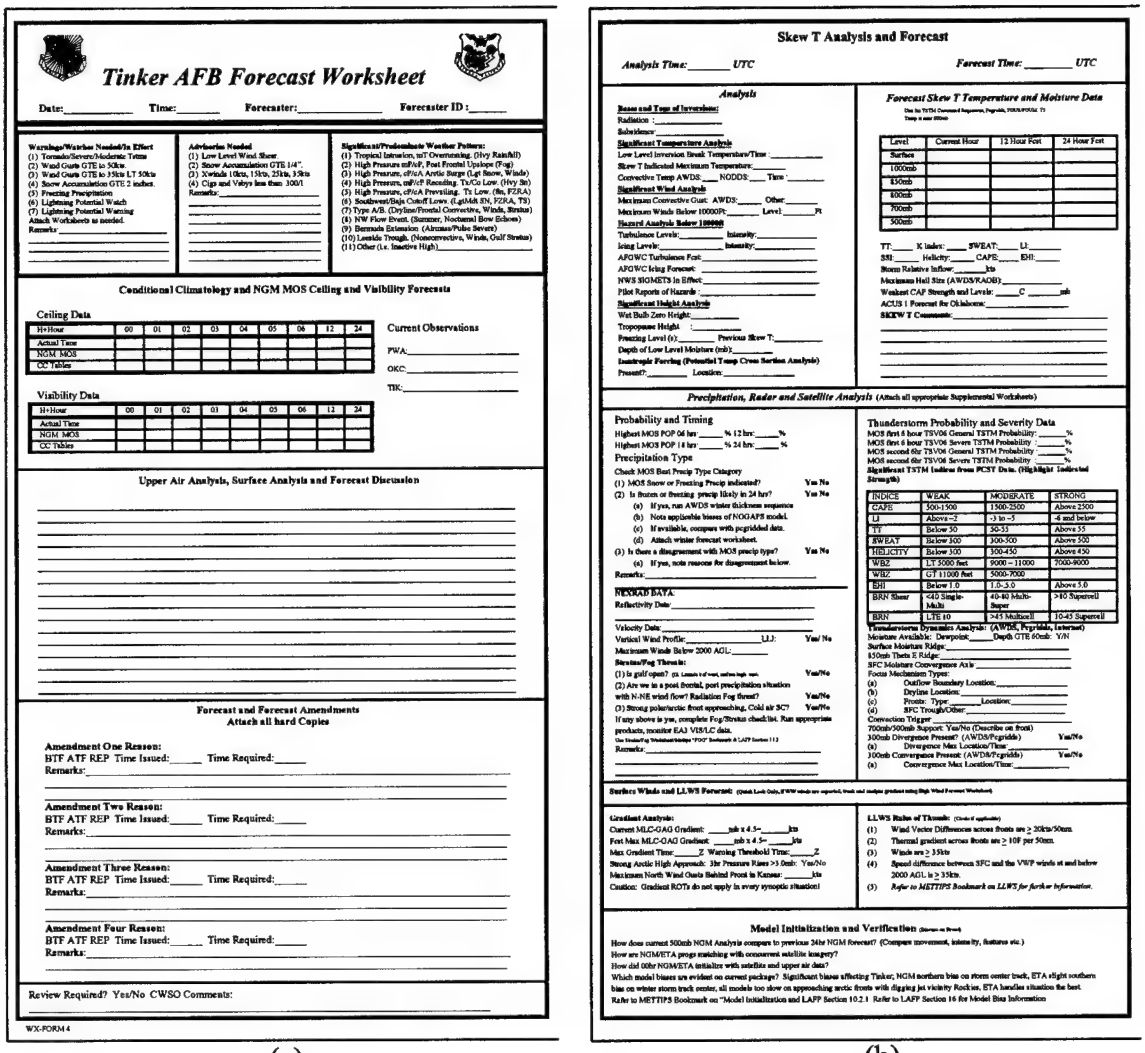
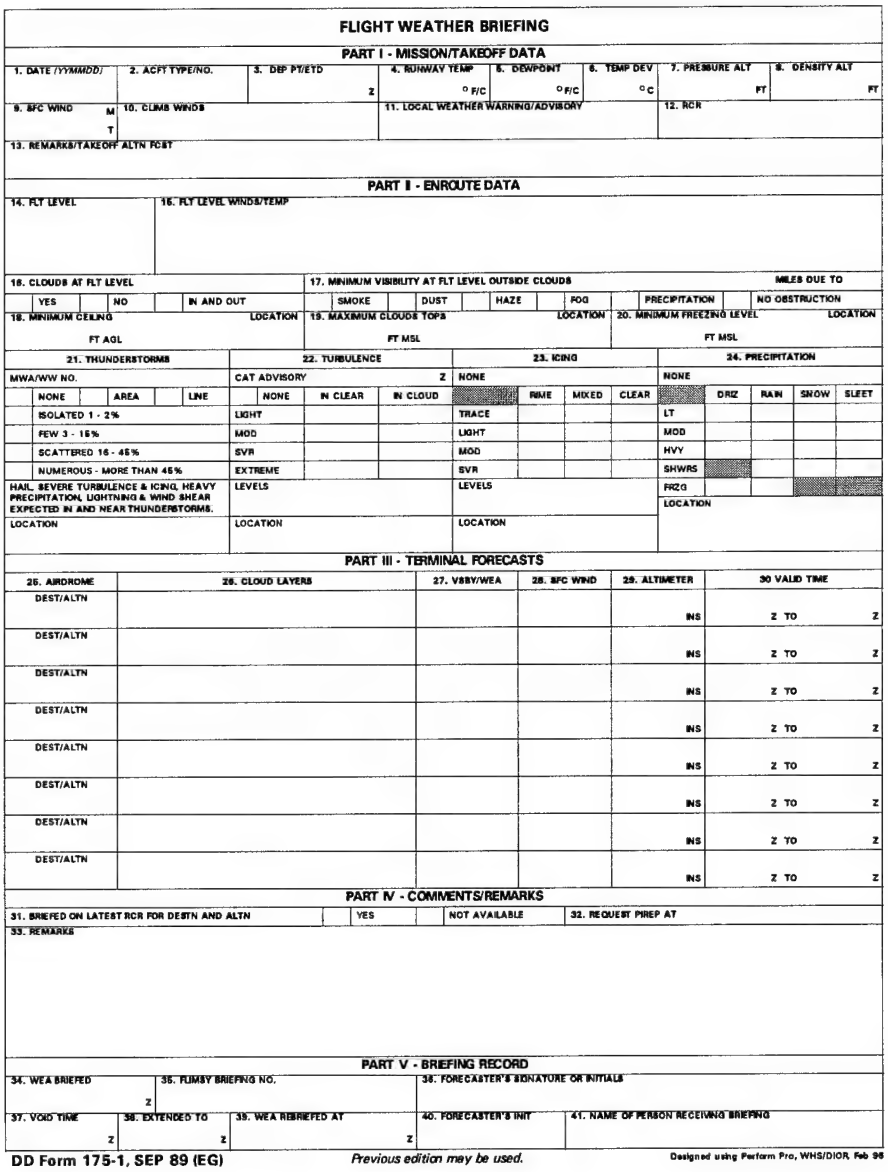


Figure 2.8: Tinker Air Force Base Terminal Aerodrome Forecast (TAF) Worksheet (a) page one and (b) page two.



FLIGHT WEATHER BRIEFING

PART I - MISSION/TAKEOFF DATA

1. DATE (YYMMDD)	2. ACFT TYPE/NO.	3. DEP PT/ETD Z	4. RUNWAY TEMP °F/C	5. DEWPOINT °F/C	6. TEMP DEV °C	7. PRESSURE ALT FT	8. DENSITY ALT FT

9. SFC WIND M T	10. CLIMB WINDS	11. LOCAL WEATHER WARNING/ADVISORY	12. RCR

13. REMARKS/TAKEOFF ALTN FCST

PART II - ENROUTE DATA

14. FLT LEVEL	15. FLT LEVEL WINDS/TEMP

16. CLOUDS AT FLT LEVEL	17. MINIMUM VISIBILITY AT FLT LEVEL OUTSIDE CLOUDS	MILES DUE TO
YES NO IN AND OUT	SMOKE DUST HAZE FOG PRECIPITATION	NO OBSTRUCTION
18. MINIMUM CEILING FT AGL LOCATION	19. MAXIMUM CLOUDS TOPS FT MSL LOCATION	20. MINIMUM FREEZING LEVEL FT MSL LOCATION

21. THUNDERSTORMS	22. TURBULENCE			23. ICING				24. PRECIPITATION				
MWA/WW NO.	CAT ADVISORY Z			NONE				NONE				
NONE AREA LINE	NONE	IN CLEAR	IN CLOUD		RIME	MIXED	CLEAR		DRIZ	RAIN	SNOW	SLEET
ISOLATED 1 - 2%	LIGHT			TRACE				LT				
FEW 3 - 15%	MOD			LIGHT				MOD				
SCATTERED 16 - 45%	SVR			MOD				HVY				
NUMEROUS - MORE THAN 45%	EXTREME			SVR				SHWRS				
HAIL, SEVERE TURBULENCE & ICING, HEAVY PRECIPITATION, LIGHTNING & WIND SHEAR EXPECTED IN AND NEAR THUNDERSTORMS.	LEVELS			LEVELS				FRZG				
								LOCATION				
LOCATION	LOCATION			LOCATION								

PART III - TERMINAL FORECASTS

25. AIRDROME	26. CLOUD LAYERS	27. VSBY/WEA	28. SFC WIND	29. ALTIMETER	30 VALID TIME
DEST/ALTN				INS	Z TO Z
DEST/ALTN				INS	Z TO Z
DEST/ALTN				INS	Z TO Z
DEST/ALTN				INS	Z TO Z
DEST/ALTN				INS	Z TO Z
DEST/ALTN				INS	Z TO Z
DEST/ALTN				INS	Z TO Z
DEST/ALTN				INS	Z TO Z

PART IV - COMMENTS/REMARKS

31. BRIEFED ON LATEST RCR FOR DESTN AND ALTN	YES	NOT AVAILABLE	32. REQUEST PIREP AT
33. REMARKS			

PART V - BRIEFING RECORD

34. WEA BRIEFED Z	35. FLIMSY BRIEFING NO.	36. FORECASTER'S SIGNATURE OR INITIALS

37. VOID TIME Z	38. EXTENDED TO Z	39. WEA REBRIEFED AT Z	40. FORECASTER'S INIT	41. NAME OF PERSON RECEIVING BRIEFING

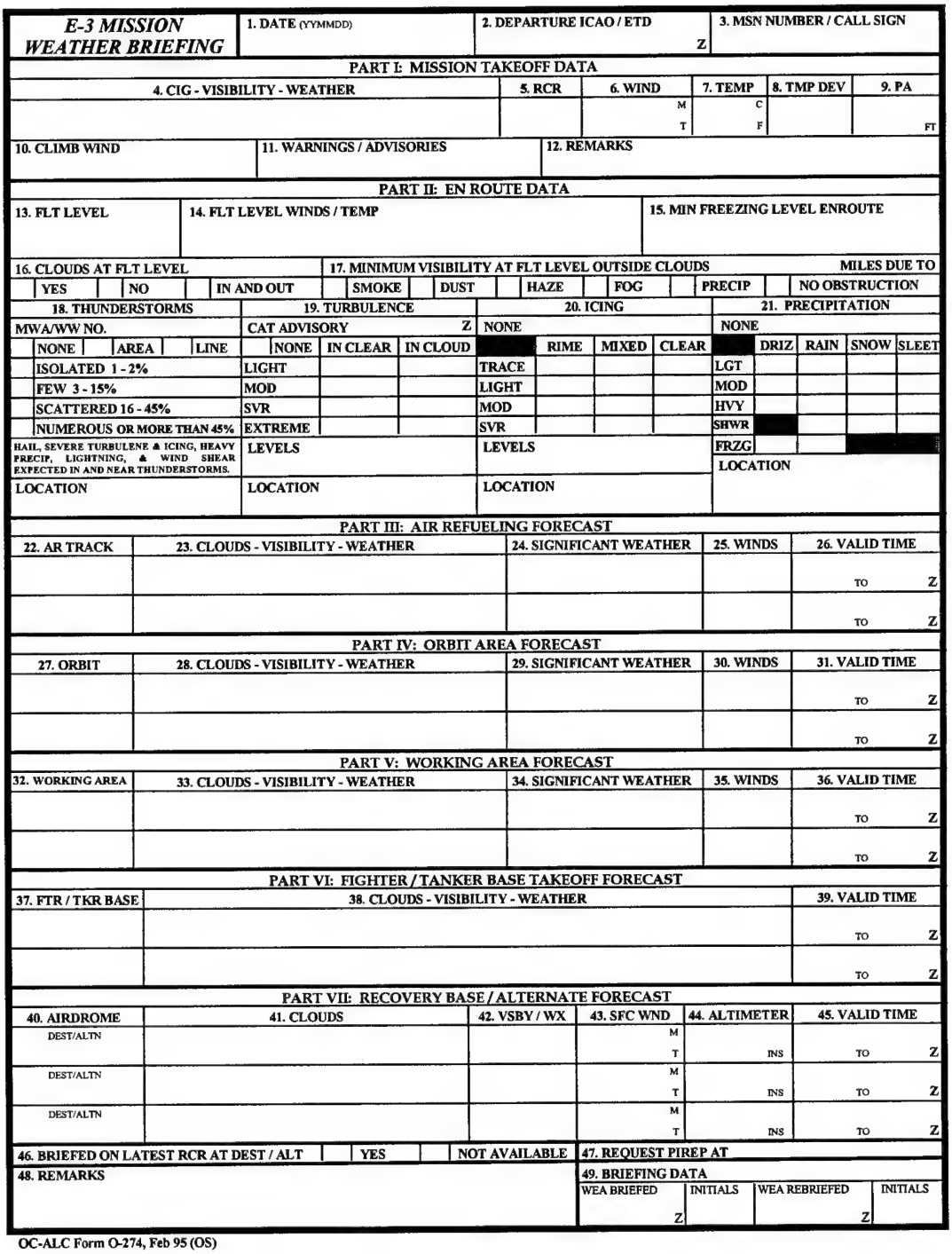
DD Form 175-1, SEP 89 (EG) *Previous edition may be used.* Designed using Perform Pro, WHS/DIOR, Feb 96

Figure 2.9: Department of Defense Form 175-1, (DD175-1) Flight Weather Briefing.

and any subsequent departures where multiple legs of the flight are expected before reaching the final destination. In some instances upwards of ten additional landing/departure locations are required due to mission restrictions, training level of the pilot, and other factors. For flights with multiple legs, the form must be updated at the appropriate weather station before the flight can continue. This ensures that any changes to in-flight and destination conditions are received. Flights of local aircraft such as the 507th and STRATCOMWINGONE require the forecaster to continuously monitor the weather throughout the planned mission duration and relay information on changes to the pilot as needed. In addition, the unpredictable nature of transient aircraft means that the forecaster must be able to quickly assess the weather at any location for updating an existing weather briefing. The final two sections are simply for documenting administrative requirements.

In the case of the AWACS, the unique nature of their mission, the extended period of operation, and the multitude of areas where they must coordinate aircraft operations all lead to a need for additional information beyond what the standard flight weather briefing provides. Tinker has modified the DD 175-1 to accommodate the extra information needed with OC-ALC Form O-274, E-3 Mission Weather Briefing, as shown in Figure 2.10. The significant additions are the inclusion of specific entries for the orbit, control, and refueling areas needed for the particular mission.

The Pilot to Metro Service (PMSV) service is provided for the handling of weather information requests from aircraft while in-flight. Pilots use radio communication with the weather station to receive arrival conditions, updates to hazardous or other significant weather conditions in the area, provide pilot reports of



E-3 MISSION WEATHER BRIEFING	1. DATE (YYMMDD)	2. DEPARTURE ICAO / ETD Z	3. MSN NUMBER / CALL SIGN

PART I: MISSION TAKEOFF DATA

4. CIG - VISIBILITY - WEATHER	5. RCR	6. WIND	7. TEMP	8. TMP DEV	9. PA
		M T	C F		FT

10. CLIMB WIND	11. WARNINGS / ADVISORIES	12. REMARKS

PART II: EN ROUTE DATA

13. FLT LEVEL	14. FLT LEVEL WINDS / TEMP	15. MIN FREEZING LEVEL ENROUTE

16. CLOUDS AT FLT LEVEL			17. MINIMUM VISIBILITY AT FLT LEVEL OUTSIDE CLOUDS						MILES DUE TO
YES	NO	IN AND OUT	SMOKE	DUST	HAZE	FOG	PRECIP	NO OBSTRUCTION	

18. THUNDERSTORMS			19. TURBULENCE			20. ICING				21. PRECIPITATION				
MWA/WW NO.			CAT ADVISORY Z			NONE				NONE				
NONE	AREA	LINE	NONE	IN CLEAR	IN CLOUD		RIME	MIXED	CLEAR		DRIZ	RAIN	SNOW	SLEET
ISOLATED 1 - 2%			LIGHT			TRACE				LGT				
FEW 3 - 15%			MOD			LIGHT				MOD				
SCATTERED 16 - 45%			SVR			MOD				HVY				
NUMEROUS OR MORE THAN 45%			EXTREME			SVR				SHWR				
HAIL, SEVERE TURBULENE & ICING, HEAVY PRECIP, LIGHTNING, & WIND SHEAR EXPECTED IN AND NEAR THUNDERSTORMS.			LEVELS			LEVELS				FRZG				
										LOCATION				
LOCATION			LOCATION			LOCATION								

PART III: AIR REFUELING FORECAST

22. AR TRACK	23. CLOUDS - VISIBILITY - WEATHER	24. SIGNIFICANT WEATHER	25. WINDS	26. VALID TIME
				TO Z
				TO Z

PART IV: ORBIT AREA FORECAST

27. ORBIT	28. CLOUDS - VISIBILITY - WEATHER	29. SIGNIFICANT WEATHER	30. WINDS	31. VALID TIME
				TO Z
				TO Z

PART V: WORKING AREA FORECAST

32. WORKING AREA	33. CLOUDS - VISIBILITY - WEATHER	34. SIGNIFICANT WEATHER	35. WINDS	36. VALID TIME
				TO Z
				TO Z

PART VI: FIGHTER / TANKER BASE TAKEOFF FORECAST

37. FTR / TKR BASE	38. CLOUDS - VISIBILITY - WEATHER	39. VALID TIME
		TO Z
		TO Z

PART VII: RECOVERY BASE / ALTERNATE FORECAST

40. AIRDROME	41. CLOUDS	42. VSBY / WX	43. SFC WND	44. ALTIMETER	45. VALID TIME
DEST/ALTN			M T	INS	TO Z
DEST/ALTN			M T	INS	TO Z
DEST/ALTN			M T	INS	TO Z

46. BRIEFED ON LATEST RCR AT DEST / ALT	YES	NOT AVAILABLE	47. REQUEST PIREP AT
48. REMARKS			49. BRIEFING DATA

WEA BRIEFED	INITIALS	WEA REBRIEFED	INITIALS
Z		Z	

OC-ALC Form O-274, Feb 95 (OS)

Figure 2.10: 552nd Air Control Wing, E-3 Flight Weather Briefing.

weather at altitude, or get other weather related information as required. The system enables contacts with aircraft to a range of approximately 200 miles at or above a flight level of 20,000 feet.

The last set of services critical to supported customers is related to the safety and protection of resources from hazardous and other severe weather. Of specific interest are those that will affect operations or pose a threat to property or life. The weather station has the responsibility to provide timely notification for any conditions affecting the operations of units located at Tinker and those units from other bases planning to conduct operations at or in the vicinity of the base. The weather station accomplishes this through a Watch/Warning/Advisory system.

A Weather Watch is designed to provide advance notice, at least three to six hours ahead of time, of weather conditions expected to impact the area. The specific criteria for issuing a watch, based on specific customer dictated or Air Force standard requirements are: tornadic activity, severe thunderstorm activity (wind speeds greater than or equal to fifty knots or hail greater than or equal to three-quarters of an inch), moderate thunderstorms (wind speeds between thirty-five and fifty knots or hail between one-quarter and three-quarters of an inch), any freezing precipitation, lightning potential, and snowfall accumulation greater than two inches within a twelve hour period.

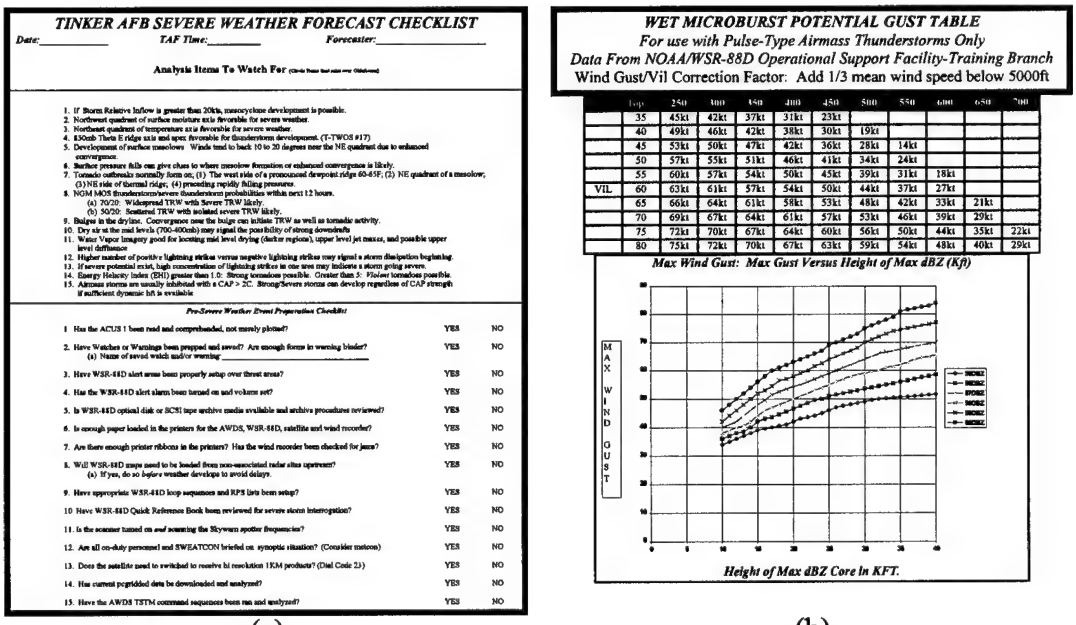
A Weather Warning is the next step in notification of hazardous weather. It provides notice of impending significant weather conditions within five nautical miles of the base. Each warning criteria also stipulates a desired lead time. This is the amount of time, in advance of the event's occurrence, that base units need to take specific precautions such as putting aircraft in hangars, clearing the airfield of personnel, or other

actions to protect resources. The specific criteria for issuing a warning are: tornadic activity; hail greater than or equal to one-quarter of an inch; two separate surface wind speed criteria: one for wind speeds between thirty-five and fifty knots and one for winds greater than or equal to fifty knots; freezing precipitation; lightning potential; and snowfall accumulation greater than two inches within a twelve hour period. The respective lead time requirements are: ten minutes, one hour, thirty minutes, one hour, one hour, when observed, and two hours. These warning conditions coincide with listed watch conditions, but can be issued without a valid watch in effect if conditions warrant.

The Advisory is primarily used to notify when specific conditions are *observed* on station. Advisories are appended to the local version of the current observation to provide details about the current conditions to the local units. Observed conditions for issuing an advisory are: a tornado, a thunderstorm within twenty-five or ten nautical miles, lightning potential, ceiling or visibility below three hundred feet or one nautical mile respectively, wind speeds across the orientation of the active runway above specific thresholds ranging from ten to thirty-five knots, and low-level wind shear. The only *forecast* advisory is for expected snowfall accumulation between one-fourth and two inches, with a desired lead time of sixty minutes. Once the event is no longer observed, the advisory is canceled.

To aid in the production of forecasts related to significantly hazardous weather, additional worksheets have been prepared. As with the TAF worksheet, these write-ups are designed to lead the forecaster through a logical thought process in determining the potential for damaging weather, assist new forecasters in assessing conditions for specified events, provide local "rules of thumb", and allow the computation and

documentation of specific actions required in the event the indicated weather occurs. Worksheets have been made to handle severe thunderstorms, microbursts and other non-convective damaging wind events, and winter time forecasting of freezing precipitation or heavy snowfall. In many cases these specialized worksheets are required if the current TAF includes various forecast elements. As an example, a severe weather worksheet, illustrated in Figure 2.11, is required if the forecast includes any mention of hail or other severe thunderstorm activity.



TINKER AFB SEVERE WEATHER FORECAST CHECKLIST

Date:_____ TAF Time:_____ Forecaster:_____

Analysis Items To Watch For

1. If Storm Relative Inflow is greater than 20kts, mesocyclone development is possible.
2. Northwest quadrant of surface moisture axis favorable for severe weather.
3. Northeast quadrant of temperature axis favorable for severe weather.
4. 850mb Theta E ridge axis and spec favorable for thunderstorm development. (T-TWOS #17)
5. Development of surface mesolows: Winds tend to back 10 to 20 degrees near the NE quadrant due to enhanced convergence.
6. Surface pressure falls can give clues to where mesolow formation or enhanced convergence is likely.
7. Tornado outbreaks normally form on; (1) The west side of a pronounced dewpoint ridge 60-65F; (2) NE quadrant of a mesolow; (3) NE side of thermal ridge; (4) preceding rapidly falling pressures.
8. NGM MOS thunderstorm/severe thunderstorm probabilities within next 12 hours.
   (a) 70/20: Widespread TRW with Severe TRW likely.
   (b) 50/20: Scattered TRW with isolated severe TRW likely.
9. Bulges in the dryline. Convergence near the bulge can initiate TRW as well as tornadic activity.
10. Dry air at the mid levels (700-600mb) may signal the possibility of strong downdrafts
11. Water Vapor imagery good for locating mid level drying (darker regions), upper level jet maxes, and possible upper level diffluence
12. Higher number of positive lightning strikes versus negative lightning strikes may signal a storm dissipation beginning.
13. If severe potential exist, high concentration of lightning strikes in one area may indicate a storm going severe.
14. Energy Helicity Index (EHI) greater than 1.0: Strong tornadoes possible. Greater than 2: Violent tornadoes possible.
15. Airmass storms are usually inhibited with a CAP > 2C. Strong/Severe storms can develop regardless of CAP strength if sufficient dynamic lift is available

Pre-Severe Weather Event Preparation Checklist

1. Has the ACUS 1 been read and comprehended, not merely plotted?	YES	NO
2. Have Watches or Warnings been prepped and saved? Are enough forms in warning binder? (a) Name of saved watch and/or warning _____	YES	NO
3. Have WSR-88D alert areas been properly setup over threat areas?	YES	NO
4. Has the WSR-88D alert alarm been turned on and volume set?	YES	NO
5. Is WSR-88D optical disk or SCSI tape archive media available and archive procedures reviewed?	YES	NO
6. Is enough paper loaded in the printers for the AWDS, WSR-88D, satellite and wind recorder?	YES	NO
7. Are there enough printer ribbons in the printers? Has the wind recorder been checked for jams?	YES	NO
8. Will WSR-88D maps need to be loaded from non-associated radar sites upstream? (a) If yes, do so *before* weather develops to avoid delays.	YES	NO
9. Have appropriate WSR-88D loop sequences and RPS lists been setup?	YES	NO
10. Have WSR-88D Quick Reference Book been reviewed for severe storm interrogation?	YES	NO
11. Is the scanner tuned on *and* scanning the Skywarn spotter frequencies?	YES	NO
12. Are all on-duty personnel and SWEATCON briefed on synoptic situation? (Consider meteos)	YES	NO
13. Does the satellite need to switched to receive hi resolution 1KM products? (Dial Code 23)	YES	NO
14. Has current regridded data be downloaded and analyzed?	YES	NO
15. Have the AWDS TSTM command sequences been run and analyzed?	YES	NO

(a)

WET MICROBURST POTENTIAL GUST TABLE

For use with Pulse-Type Airmass Thunderstorms Only

Data From NOAA/WSR-88D Operational Support Facility-Training Branch

Wind Gust/Vil Correction Factor: Add 1/3 mean wind speed below 5000ft

Top	250	300	350	400	450	500	550	600	650	700
35	45kt	42kt	37kt	31kt	23kt					
40	49kt	46kt	42kt	38kt	30kt	19kt				
45	53kt	50kt	47kt	42kt	36kt	28kt	14kt			
50	57kt	55kt	51kt	46kt	41kt	34kt	24kt			
55	60kt	57kt	54kt	50kt	45kt	39kt	31kt	18kt		
VIL 60	63kt	61kt	57kt	54kt	50kt	44kt	37kt	27kt		
65	66kt	64kt	61kt	58kt	53kt	48kt	42kt	33kt	21kt	
70	69kt	67kt	64kt	61kt	57kt	53kt	46kt	39kt	29kt	
75	72kt	70kt	67kt	64kt	60kt	56kt	50kt	44kt	35kt	22kt
80	75kt	72kt	70kt	67kt	63kt	59kt	54kt	48kt	40kt	29kt

Max Wind Gust: Max Gust Versus Height of Max dBZ (Kft)

MAX WIND GUST

Height of Max dBZ Core in KFT.

(b)

Figure 2.11: Tinker Air Force Base Severe Weather Worksheet page (a) one and (b) two.

Section 2.2 Project COMET-Tinker

In an effort to evaluate the ability to apply fine scale numerical weather prediction (NWP) to Air Force Base Weather Station forecasting needs, a three year project nicknamed COMET-Tinker was initiated in 1996. The project is a joint venture between the University of Oklahoma's (OU) School of Meteorology, the Cooperative Program for Operational Meteorology Education and Training (COMET) Outreach Program, and the Air Force Weather Agency (AFWA) [formerly the Air Weather Service (AWS)] of the United States Air Force. The desire of the project was to center forecasting activities around the needs of Tinker AFB, located within thirty minutes of the OU campus. The goals of the project are: (1) to demonstrate the capability to provide real-time, storm-scale numerical weather prediction guidance to an Air Force Base Weather Station; and (2) to assess the value of these high-resolution forecasts and the suite of requested products derived from them from an operational perspective of an AWS forecaster (Carr et al. 1996).

Using the understanding of the normal operations of the weather station at Tinker AFB outlined earlier, and with close coordination between Air Force forecasting personnel at Tinker and project personnel at OU, a list of desired forecast products was established. In addition to the standard meteorological fields routinely used in analyzing model output (e.g. surface and constant pressure level temperatures, winds, and moisture indications) the developed forecast product "suite" included indications of specific weather phenomena of interest to duty forecaster and the operational units supported. Table 2.3 shows the initial list of AIVs that were requested to be incorporated into

Product	Details / Examples
Icing *	Type: Rime, Clear, or Mixed Intensity: Trace, Light, Moderate, or Severe
Turbulence *	Type: In Cloud, Clear Air, and Continuous/Intermittent Intensity: Light, Moderate, or Severe
Visibility *	Restrictions: Fog, Haze, Rain, or Dust Cutoff Value: 1/2, 1, 2, 3, and 7nautical miles
Lowest Ceiling	Cutoffs: 500, 1000, 2000, and 3000 feet
Standard Stability Indices *	CAPE, Lifted Index, Total Totals , K Index,
Hail Potential/Size *	Low, Moderate, or High Cutoffs: 1/4, 1/2, and 1 inch
Wind Gusts / Downburst Potential *	Cutoffs: +10, +20 knots, and +30 knots, etc
Lightning Potential *	Low, Moderate, or High
Low Level Wind Shear *	Yes/No
Precipitation Type *	Rain, Snow, Sleet, Freezing Rain, or Mixed
Precipitation Accumulation *	Hourly rates, 3 hour amounts, 6 hour amounts
Cloud Type *	Standard cloud types
Storm Tracking Information	Cell location, movement, and properties
Radar Reflectivity Echo Top *	Height if highest reflectivity forecast
Comfort Indices *	Heat Stress Index and Wind Chill
Surface Meteograms *	Temperature, Dewpoint, Wind Direction and Speed, Precipitation Amount, and Sea Level Pressure or Altimeter Setting
Profiler Meteograms	Winds Direction and Speed
Runway Crosswinds *	Cutoffs: 10, 15, 25, and 35 knots
Runway Condition	Wet, Dry, Ice, or Mixed
Isentropic Plots *	Temperature, Relative Humidity, Heights, Wind Barbs, Montgomery Stream Function
Change Indications	1 and 3 hour; Temperature, Dewpoint, Height
Altimeter Setting *	Horizontal spatial plot of altimeter settings
Mean Fields	1000-500mb Thickness, SFC-850 Relative Humidity, etc.

Table 2.3: Initial forecast product and aviation impact variable (AIV) list. (* indicates a product has been incorporated, in some manner, into the real-time production routines)

the suite of products. Given the extensive data sources available to the Air Force forecaster, these AIVs were designed to eliminate the need of the forecaster to perform a full "analysis" on yet another source of NWP information. From the list we see that many of the desired products also correspond directly to specific criteria governing the services provided by the weather station. Forecasts of icing and turbulence can be directly incorporated into the terminal forecast and flight weather briefings, while hail and wind speed (gust) indications can be used to determine the need for specific watch or warning preparation. In this way information can be presented to the forecaster in a value-added context, without overloading the forecaster even further.

The initial setup of the project was meant to leverage off the activities of CAPS for production of forecast information. CAPS has conducted various real-time forecasting periods using its Advanced Regional Prediction System (ARPS; Xue et al. 1995) storm-scale prediction model for research and operational forecasting interests. In both 1995 and 1996 CAPS ran daily model forecasts for portions of the Southern Plains, centered on Oklahoma, during the springtime severe weather season. Forecasts were performed using the resources at the Pittsburgh Supercomputing Center (PSC). Model runs ranged in resolution down to three kilometers, with six to nine hour forecast windows during the anticipated afternoon convective periods (e.g. 15 to 06 UTC). Forecasting personnel from the National Weather Service Forecast Office (NWSFO) in Norman, OK, the National Severe Storms Laboratory (NSSL), the Storm Prediction Center (SPC), and researchers at OU have all used the forecasts for various activities. A goal of CAPS for 1997 and beyond was to make year-round operational forecasts available. This became possible when CAPS began a three-year joint venture with

American Airlines called Project Hub-CAPS (Carpenter et al. 1997). With goals similar to those of COMET-Tinker, Hub-CAPS' focus is on the production and integration of storm-scale NWP into weather service operations at American Airlines. Hub-CAPS conducted a Spring Operational Forecast Period in May of 1997, as well as a Winter Operational Forecast Period from December 1997 to January 1998, in which daily 9 km forecasts were accomplished. In addition, Hub-CAPS began producing ongoing twenty-seven kilometer resolution forecasts, initiated every six hours, in late November of 1997, continuing to the present. The operational forecasts made by Hub-CAPS are the primary source of data used in creating forecast products for the COMET-Tinker efforts. Hub-CAPS also generates hourly observational analyses, using the ARPS Data Analysis System (ADAS; Brewster et al. 1994), which are used by COMET-Tinker for real-time analysis based assessment of the same AIVs produced for forecast. The forecast data set used for this study consisted of nine kilometer model runs made by Hub-CAPS between 23 December 1997 and 31 January 1998. Details regarding the model configuration and setup are in Section 2.3.

To determine individual AIV forecasts, numerical computations are performed on the model output generated by Hub-CAPS. In many instances this simply eliminates the need of the forecaster to do manual calculations by automating procedures already in place. Calculation of standard indices such as Convective Available Potential Energy (CAPE) and Lifted Index (LI), and determination of runway crosswind components can easily be incorporated into post-processing routines. For other AIVs, the incorporation of empirical and other research-based schemes were introduced by the COMET-Tinker team. Post-processing of forecast data is accomplished by the COMET-Tinker project

using a SUN UltraSparc 1, model 140, workstation located on the OU campus. Existing graphics software developed by CAPS called ARPSPLT was used as the starting point for all post processing development. Written in Fortran-77, code for each additional AIV (sample image shown in Figure 2.12) was added to the ARPSPLT graphics package as it was developed, with the overall forecast generation process controlled by UNIX shell command scripts.

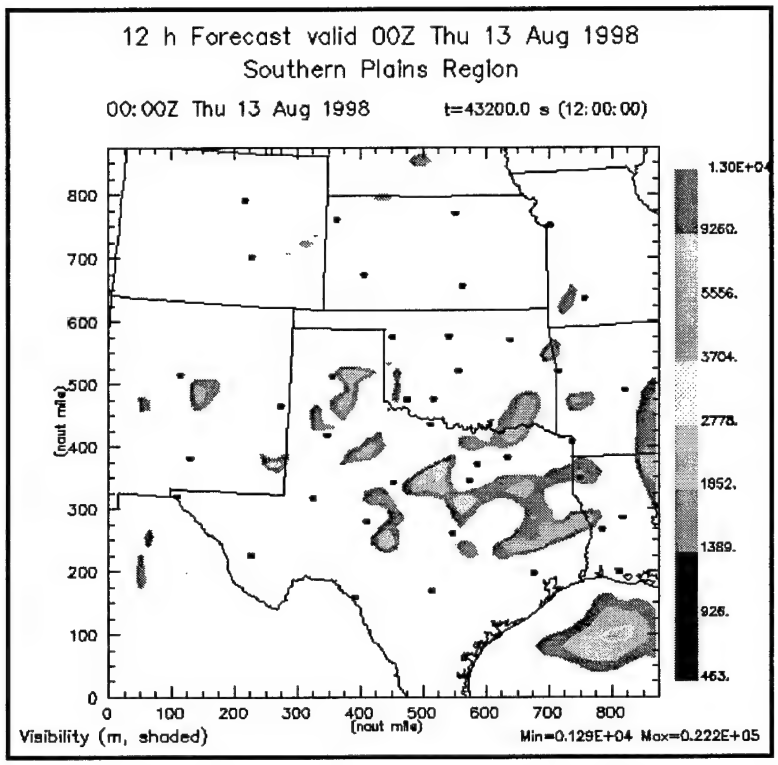


Figure 2.12: Sample forecast product image of surface visibility.

The goals of the COMET-Tinker project involved integrating the storm-scale forecast products into the daily routines of forecasters at Tinker and monitoring the usefulness of the information provided. To make information available via a graphical interface that would require minimal training time for the forecasters to operate, the World Wide Web (WWW) was utilized to disseminate products to the field. Most

forecasting personnel are familiar with accessing and browsing (using commercial software such as Netscape Navigator) the WWW from computers in the BWS. No specialized training and spin–up time was needed that would have been required if the project used a separate data display system. Another factor for using the WWW was the ability for more than one person to access the forecast information at the same time or from sites outside the BWS itself. A web server (http://comet-tik.ou.edu) was established on the same computer that does the post processing. Figure 2.13 shows the COMET-Tinker WWW server "Home" page.

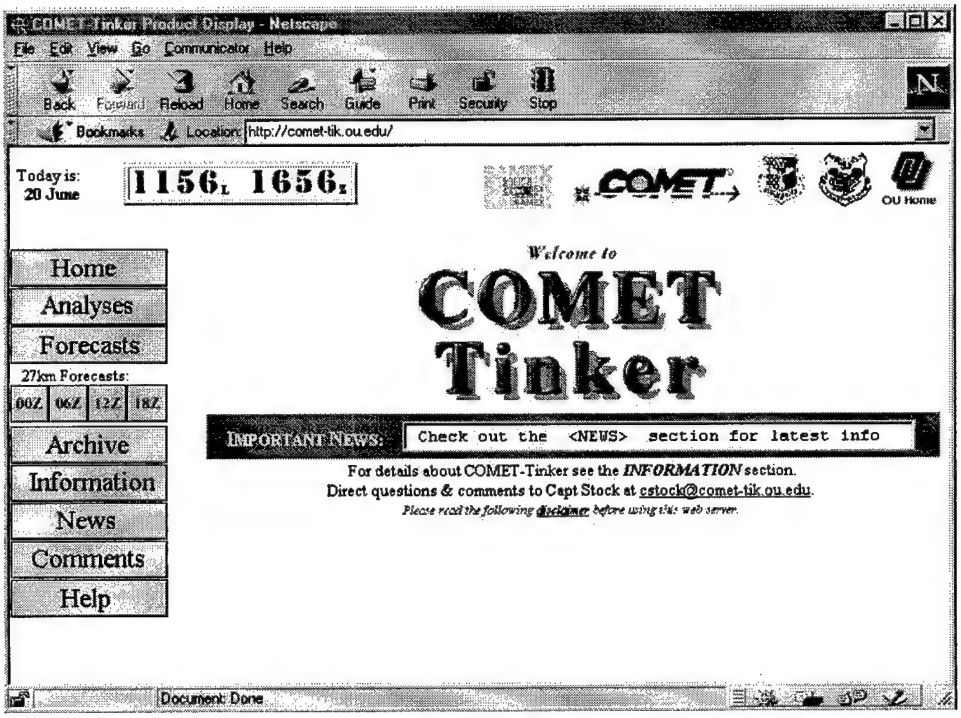


Figure 2.13: Project COMET-Tinker World Wide Web Home Page

The design emphasis of the web page interface is on the ability to easily examine items of interest. The server is broken down into the following sections: forecast and analysis products areas; real-time feedback capabilities via electronic mail (e-mail); an extensive HELP section to answer questions regarding the available products and the

project as a whole; and archive/information sections to access past forecasts, verification information for post analysis and review, and other project information. The individual server sections are displayed in a separate frame area so that the forecaster can go from section to section easily using the "sections tabs" on the left side of the main display. The current date and time, both local and Greenwich Mean Time (GMT or UTC) are displayed at the top to allow forecasters to easily judge what products should be available and which ones may be out of date.

The forecast and analysis web pages (layouts shown Figure 2.14(a) and (b) respectively) are set up with hypertext-links (links) to the actual graphical images for the valid date/time specified at the top of the product page (not visible). Products are listed in tabular format with different product types such as surface products, upper air fields on constant pressure levels, forecast soundings, upper air AIV products at various aircraft operating altitudes (flight levels), and vertical cross-sections separated for convenience. A set of navigation links is located at the top of the product display to enable access of each table section from the screen directly. In the case of forecasts, there will be multiple valid times of interest, one for each forecast hour in the run. Thus the forecast pages also include a second list of each forecast valid time, UTC time convention, listed above the section links. During real-time operations the listing will include only those hours where the post processing generation of the forecast products has been accomplished (e.g. after processing the second hour forecast products the list will have three entries for the zero, one, and two hour forecasts respectively). Once the full forecast has been processed the list will show the complete set of forecast valid times for that particular forecast run.

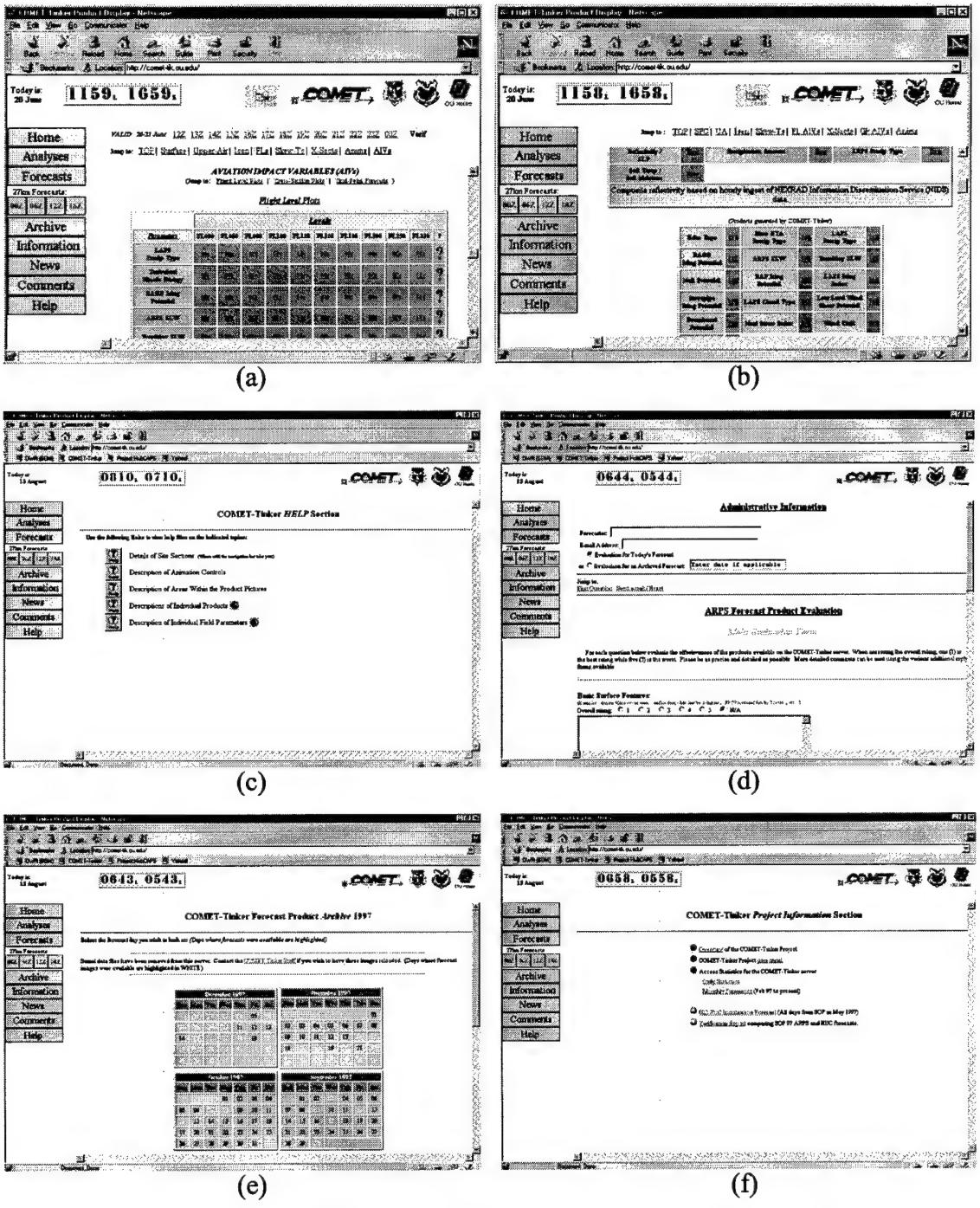


Figure 2.14: Project COMET-Tinker WWW Section pages: (a) forecast, (b) analysis, (c) individual AIV product details, (d) comments/reply, (e) product archive, (f) project information web pages.

Since the specialized AIV products use computational algorithms for determination, a means to describe how these computations were performed was needed. In this way the forecasters could easily see how individual forecast elements were derived and use its methodology to assess the usefulness and applicability to their specific forecast situation. Instead of requiring the forecasters to read the large volume of literature and other printed specifics behind each algorithm, condensed versions were integrated into a "Help" section. Various help sub-sections (see Figure 2.14(c)) give information on what forecast products are available and grouped on the same image, how to read the layout of the forecast image, and other details. On the forecast pages there are separate question mark links, beside the image display links, which are used to go directly to the help information for that product. A separate web browser screen, illustrated in Figure 2.15, allows the forecaster to read about that product and others, if desired, without going back and forth between the forecast pages and help sections.

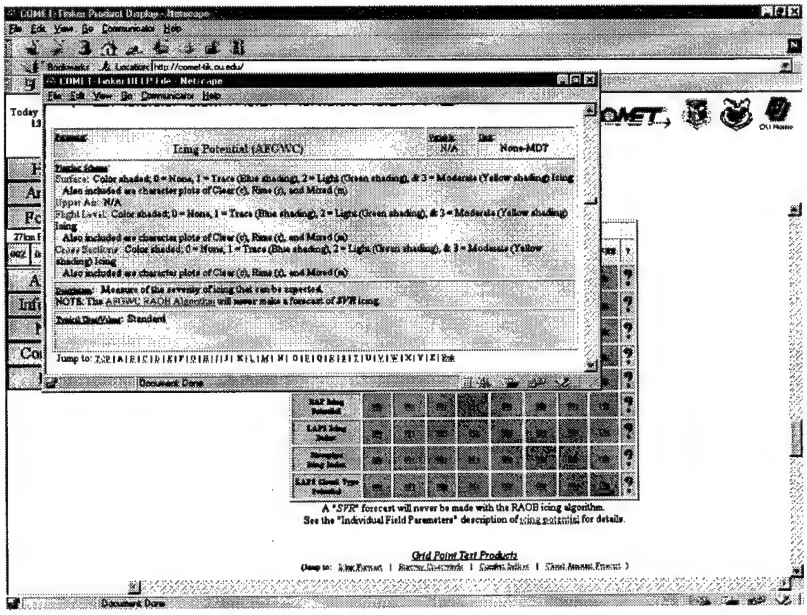


Figure 2.15: Project COMET-Tinker product "Help" information integrated within the forecast web page.

The second goal of the COMET-Tinker project specifically identifies the evaluation of the forecast information as it pertains to the operations at Tinker. In the traditional manner, verification against observational data over a full three-dimensional forecast volume is needed. To this end, verification data is archived and stored on the server for post-analysis use. Surface observations from the Oklahoma Mesonet and routine surface reporting stations, upper air soundings, and wind profiler network data are saved to do point by point verifications. In addition, constant pressure level information from the Rapid Update Cycle II (RUCII) and the ETA model are extracted to do a field (grid point) comparison of standard model output (e.g. temperature and winds). Daily and monthly averaged verification plots are also created, with the results placed on the COMET-Tinker server for access via the WWW.

It is also desirable to get feedback directly from the forecasting personnel about how the forecasts are being used, what products are being looked at, which ones are helpful, and where questions exist about what is on the system. To accomplish this, a "Comments" section of the server, illustrated in Figure 2.14(d), is used to solicit user feedback. Specific question blocks can be completed, as well using standard web form "radio button" selections. As a final check on the activity of the system, commercially developed software, Accesswatch, is used to monitor and compile server access statistics to see what product images and server web pages are being examined.

The last three areas of the server are the "Archive" and "Information" sections, Figure 2.14(e) and (f) respectively, and the "News" section (not shown). The "Archive" sections provide access to previous forecast and verification information. The "Information" sections lists administrative details about the COMET-Tinker project,

other general information, and provides the links to the verification and evaluation information that was collected. Finally the "News" section enables important status and update information to be seen from a single location.

Section 2.3 WOP97 Specifics

In order to evaluate the quality and usefulness of fine scale numerical weather prediction in the BWS setting, real-time operational forecasts must be made on a daily basis. This enables the duty forecaster to become familiar with the format and type of forecast products available, allows a routine subjective assessment of forecast ability to be made, and provides a set of forecasts on which to perform a statistical comparison against observational data. A winter operational forecasting period was conducted by CAPS through their joint effort with American Airlines during December 1997 and January 1998 in which nine-hour forecasts, at a horizontal grid spacing of 9 km, were produced using ARPS. During the early part of this period, adjustments were made to the model configuration to incorporate ongoing changes to the ARPS code, modify initialization and boundary condition requirements, and alter the starting time of the forecast runs. During the five week period from 23 December to 31 January the basic model configuration was essentially frozen with a total of thirty-five daily forecasts completed. For reference in this work, this set of forecasts is referred to as the Winter Operational Period 1997 (WOP97) and provides the forecast information used for the verification results.

The forecasts produced by Project Hub-CAPS during WOP97 were generated using the ARPS forecast model. ARPS is a non-hydrostatic, fully compressible forecast

model that utilizes a generalized terrain-following coordinate system (Xue et al. 1995). The grid spacing is uniform in the horizontal and can incorporate stretching in the vertical. During WOP97 the model was operated at two forecast resolutions, both centered near Brownwood, TX, (31 48N, 098 57W) as indicated in Figure 2.16. Forecasts for this outer domain, roughly 1620 km on a side (over 2,500,000 km^2) with a

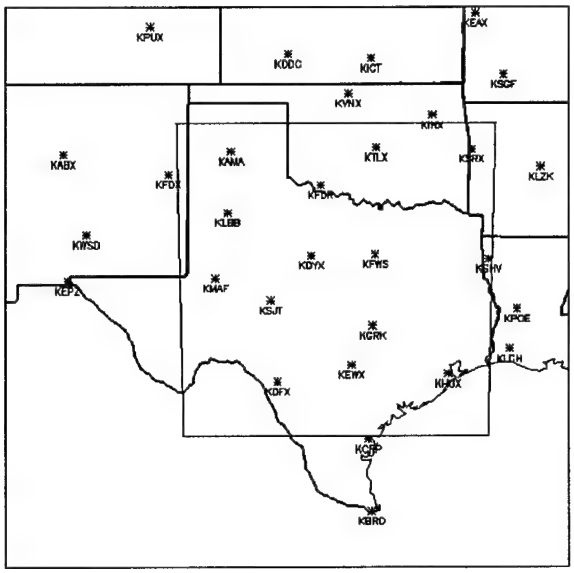


Figure 2.16: Winter Operational Period 1997 (WOP97) forecast domain. 27 km forecast domain covers entire area, 9 km forecast domain indicated by inner box. Locations of NEXRAD radar sites annotated by *.

horizontal grid spacing of 27 km, were made every six hours, starting at 00, 06, 12, and 18 UTC, out to the twelve-hour forecast point. Forecast output fields were created at hourly intervals to be used in generating the forecast products for display via the WWW, and to provide the initialization and boundary conditions for a nested inner forecast domain. The inner forecast domain (boxed region in Figure 2.16) covered 900 km on a side (810,000 km^2) at a resolution of 9 km. Both domains were run with 39 vertical levels on a stretched grid with an average spacing of 500 meters and a minimum spacing (near the ground) of 20 meters. The one hour forecast from the 12 UTC outer domain

was used as the background field initializing the 9 km forecast. Forecast processing was started at 1330 UTC and was performed by the Environmental Computing Applications System Cray J90 operated within the OU College of Geosciences. Normal processing time for the entire forecast was about two and a half hours. Forecast products were produced as the hourly forecast fields became available. The first set of products (00hr forecast) were normally available by 13 UTC with the final forecast hour products completed by 17 UTC. During processing the typical "break even" point where forecast products were available before the valid time was at the two hour (15 UTC) forecast point. A general timeline is illustrated in Figure 2.17.

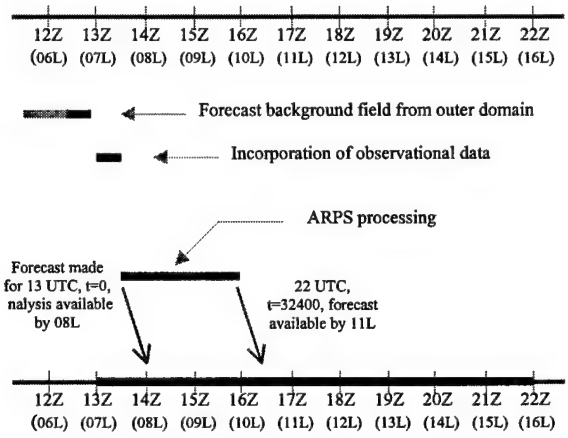
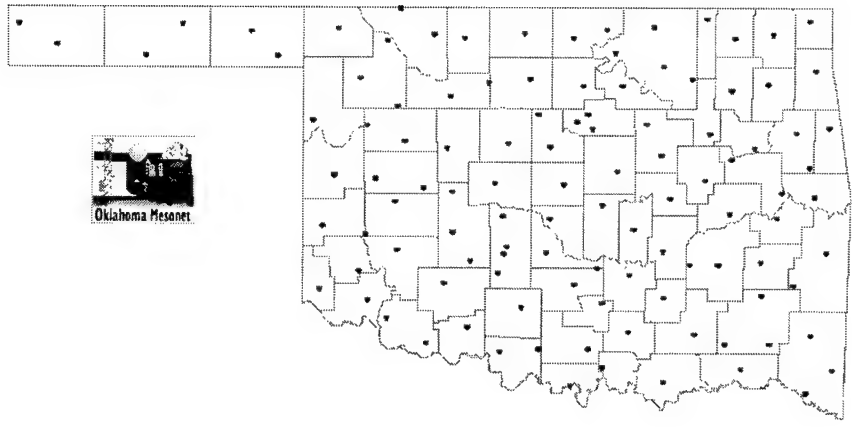
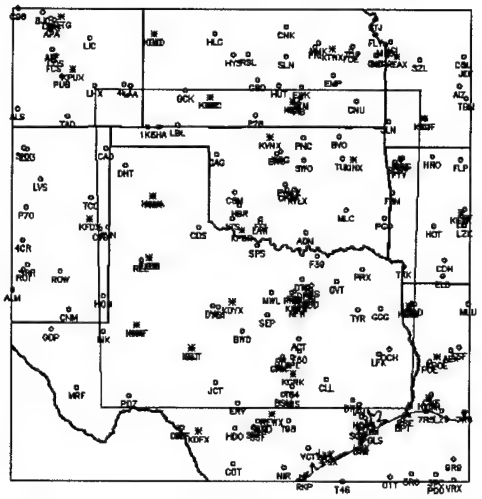


Figure 2.17: Winter Operational Period 1997 (WOP97) forecast production timeline. (L refers to local Central Standard Time)

The inner domain was initialized using the one hour forecast from the 12 UTC outer domain forecast. In addition, observational data valid at 13 UTC were incorporated into the initialized field from a variety of sources. The primary sources were observations from routine surface observing stations and the Oklahoma Mesonet (Figure 2.18), and base reflectivity NEXRAD radar observations provided by the NEXRAD



(a)



(b)

Figure 2.18: Surface observing stations within the (a) Oklahoma Mesonet and (b) surface airways observing network in the Southern Plains region.

Information Dissemination Service (NIDS). Incorporation of surface observations was performed by the ARPS Data Assimilation System (ADAS) which uses a Bratseth (1986) interactive successive correction scheme to adjust the wind components, pressure, potential temperature, and moisture parameters of the background field based on the observational data available. The scheme takes into account the observing system used, the spatial distributions of the observations, and errors inherent in the data. The NIDS data are comprised of one degree by one kilometer resolution base reflectivity values at four elevation angles (0.5, 1.45, 2.3, and 3.35 degrees). Radar locations available within the two forecast domains are shown in Figure 2.16. Interpolation of the radar reflectivity to the ARPS grid points was performed and used to adjust moisture variables within the three dimensional background field.

Version 4.3.1 of the ARPS model was used as the base code for WOP97 and incorporated many of the "high end" options available within the model code. Specifically, the model configuration included: sub-grid scale turbulent mixing parameterization and explicit treatment of vertical mixing, the Lin-Tao ice microphysics parameterization scheme (discussed in more detail in the formulation of the icing forecast algorithms), parameterizations of both shortwave and longwave radiation processes in the atmosphere, and stability-dependent calculations of surface fluxes. Readers are referred to the ARPS User's Guide (Xue et al. 1995) for additional information and specifics regarding ARPS parameterizations and overall model construction.

The weather during the entire forecast period was typical for the Southern Plains in late December and January. Four periods with synoptic scale weather cyclones (25-26 Dec, 4-8 Jan, 12-14 Jan, and 24-26 Jan) generated precipitation (both snow and rain) in

various portions of the forecast domain on the aforementioned days. Thunderstorm activity was mainly isolated in central Texas. Tinker AFB saw thunderstorms and over one-and-a-half inches of rain on 4 January and periods of freezing drizzle on the 12th and 14th (mainly during times outside the forecast window).

It is of interest to remember that the generation of forecasts during this period was made with the operational forecasting needs of American Airlines and Tinker AFB in mind. From a research standpoint, some aspects of the forecast configuration were not optimized effectively. The forecast period (13 to 22 UTC) did not overlap either upper-air data collection time (00 or 12 UTC) hence rigorous upper-air verification could not be performed. The period was also insufficient to capture significant weather events (predominantly thunderstorm activity) which might develop after the forecast period ended, or overnight precipitation events when transition from rain to snow is possible. The start time of the model was set to 13 UTC primarily to provide forecast information during peak flying operations and to facilitate nesting of the higher resolution forecast using the 27 km forecasts. Finally, the length of the forecast, the grid spacing, and domain size were limited by the computing resources available on the Cray J90.

Section 2.4 Verification Data Sources

To perform a statistical verification on the meteorological quantities produced by the ARPS forecast from WOP97, a variety of surface and upper atmospheric observations are required. Since the forecasts did not cross a 00 or 12 UTC time period, where the use of standard rawinsonde soundings or an objectively analyzed upper air analysis would be possible, a "substitute" upper air analysis was needed to evaluate the model forecasts on

constant pressure levels. Thus the initial analysis fields from the Rapid Update Cycle (RUC) forecasts, which was run every three hours, was used as a proxy. RUC analysis fields were available at 15, 18, and 21 UTC to verify the 02, 05, and 08 hours forecasts respectively. In addition, direct comparisons with hourly surface observation from regular reporting stations, data collected by the Oklahoma Mesonet, and wind profiler stations located in the forecast domain were performed. The following is a brief description of the sources of verification data used in this study.

Recent increases in automated reporting of weather information from commercial aircraft, and the implementation of a demonstration network of wind profilers, have allowed the incorporation of additional upper atmospheric information into analysis and numerical weather prediction systems available to the operational forecasting community. Forecasts can now be initiated at times other than the standard 00 and 12 UTC start times, when rawinsonde data are available, to provide additional forecast guidance throughout the day. To accomplish this, NOAA's Forecast Systems Laboratory (FSL) developed the Mesoscale Analysis and Prediction System (MAPS), a data assimilation, analysis, and forecast system that could be updated at a frequency of every three hours. In 1994, a version of the MAPS software, named the Rapid Update Cycle (RUC), was incorporated into the real-time operations of the National Centers for Environmental Prediction (NCEP – formerly the National Meteorological Center (NMC)) (Benjamin et al. 1994).

The RUC uses a hybrid isentropic-sigma vertical coordinate system where the majority of layers are represented by *isentropic* surfaces (defined as constant virtual potential temperature surfaces) except near the ground where a terrain-following (sigma)

coordinate system is used. The isentropic levels range from 224 to 410 Kelvin (K) with surfaces following the terrain as they approach the ground as shown in Figure 2.19.

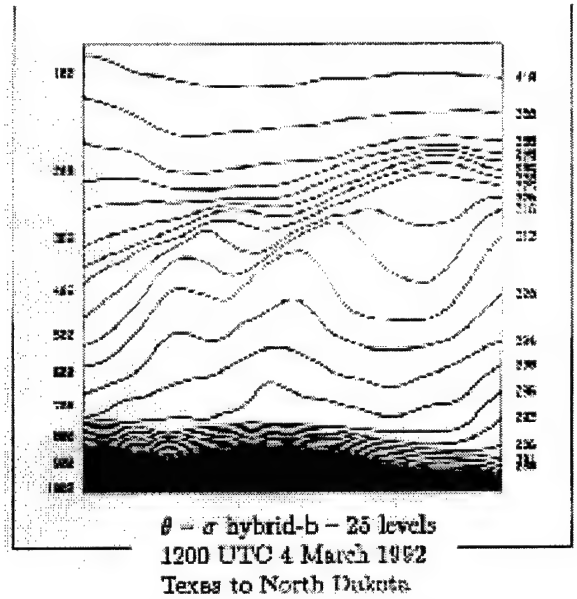


Figure 2.19: Cross section of the Rapid Update Cycle (RUC) hybrid-b isentropic-sigma coordinate system surfaces from North Texas (left) to North Dakota (right) (after Benjamin et al. 1994)

The horizontal grid of the RUC during WOP97 was 81 by 62 points with a grid spacing of 60 km at 40° N latitude. The forecast domain covers the lower forty-eight states and adjacent areas of Canada, Mexico, the Atlantic Ocean, and the Pacific Ocean. The domain is illustrated in Figure 2.20.

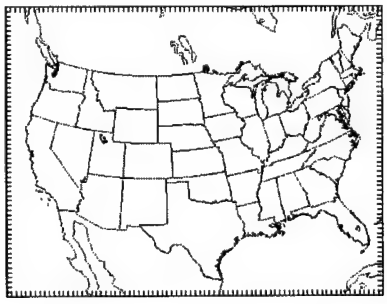


Figure 2.20: Horizontal domain of the Rapid Update Cycle (RUC) (after Benjamin et al. 1994)

To generate the initialization field for the next model run, the previous three hour RUC forecast is used as a background field. Data from wind profilers, aircraft reports, rawinsonde reports (if available), and surface stations are ingested with a maximum cutoff time for ingest of one hour and twenty five minutes (thirty-five minutes for surface reports). Routine quality control eliminates known systematic errors, gross reporting errors, and includes a horizontal consistency check. The resulting analysis provides six variables analyzed to the hybrid coordinate system. These include pressure, Montgomery stream function, virtual potential temperature, condensation pressure, and horizontal wind components.

To perform the verification against observed conditions the data sources available for comparison with the model forecast output were routine hourly surface observations, reports from the Oklahoma Mesonet, and data collected by the NOAA Profiler Network. In addition, voice transmitted Pilot Reports (PIREPS) were collected to perform verification between those reports and the model derived icing forecasts generated. Data from these four sources were collected from the routine archive capabilities of the OU School of Meteorology, the National Severe Storms Laboratory (NSSL), the Oklahoma Climatological Survey (OCS), and the real-time data feed provided by Project Hub-CAPS.

The forecasting requirements of the COMET-Tinker project involved the generation of AIV products such as runway crosswinds, heat stress and wind chill indices, and forecast soundings for specific locations of interest to Tinker AFB. The operational surface stations that were used in providing forecast information to the duty

forecaster are shown in Figure 2.21. To limit the verification performed to stations of particular interest to Tinker, surface verification data encompassed only these stations (versus all locations shown in Figure 2.18). Observations verified included the ambient air temperature, dew point temperature, wind speed and direction, and station pressure.

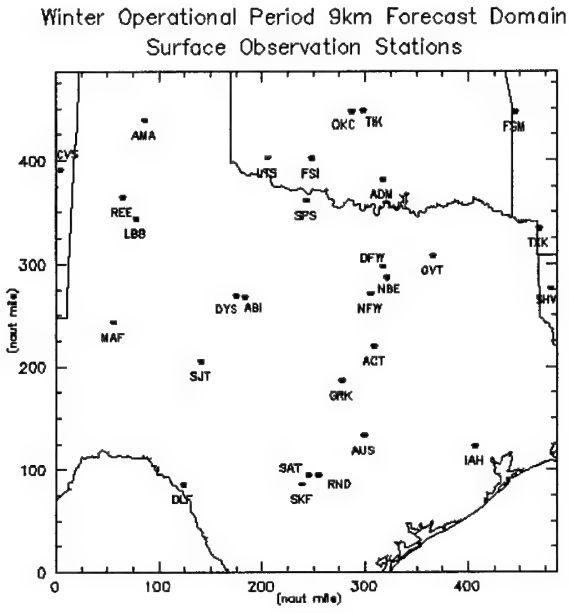


Figure 2.21: Surface observation reporting stations (listed by standard 3-digit identifier) within the forecast domain used for verification purposes.

The second source of surface data used were observations collected by the Oklahoma Mesonet (Brock et al. 1995). In the mid 1980's scientists from the University of Oklahoma and Oklahoma State University recognized the need for a fine scale surface observing network for use in agricultural, meteorological, and hydrological data gathering. The result was the creation of the Oklahoma Mesonet, which now includes 114 sites (refer to Figure 2.18(a)) across the state of Oklahoma. Mesonet instrumentation provides two to fifteen minute interval measurements of air and soil temperatures, relative humidity, barometric pressure, leaf wetness, solar radiation,

rainfall, and soil moisture. Figure 2.22 shows the Mesonet stations that were located in the forecast domain and were included in the verification. (Refer to Figure 2.21 to see the portion of Oklahoma in the forecast domain.)

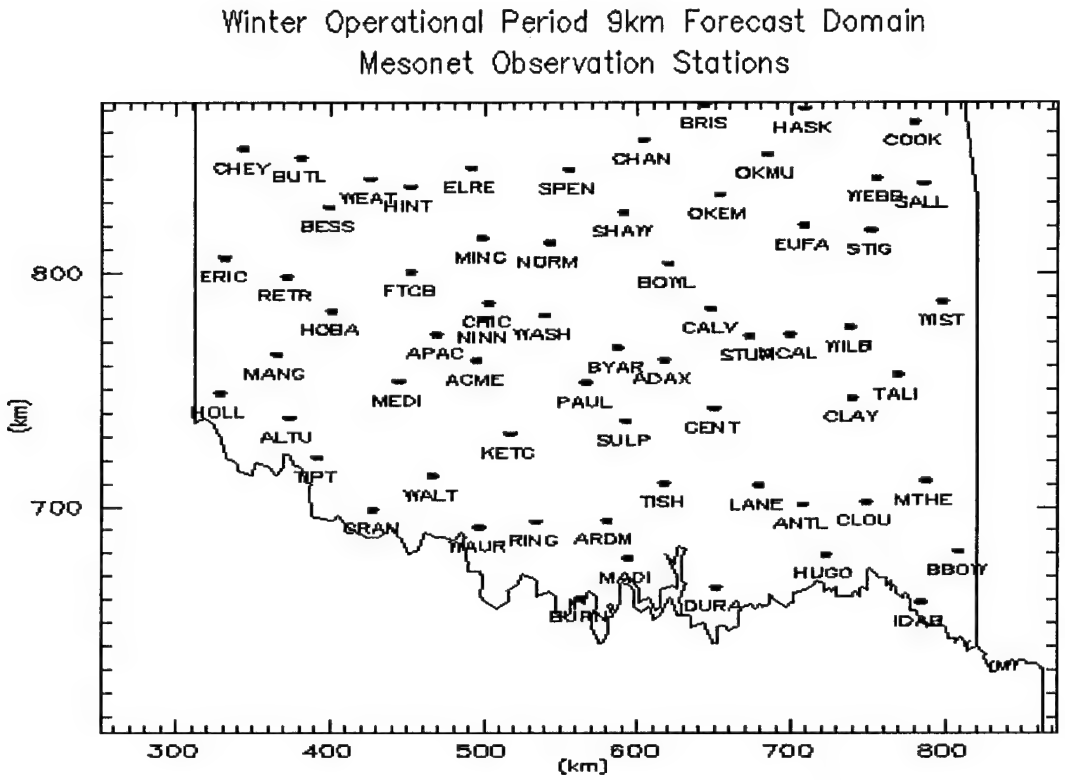


Figure 2.22: Oklahoma Mesonet stations located within the WOP97 forecast domain.

The third source of verifying data comes from the NOAA wind profiler network. Figure 2.23 shows the sites in the United States where the wind profilers are located. The wind profiler instruments detect small fluctuations in atmospheric density, caused by the turbulent mixing of air with different temperature and moisture content. The radial wind components are derived from the frequency shift of the returned energy signal from refraction regions. Measurements of the radial wind components are then converted to Cartesian wind components at six minute intervals beginning 500 meters above the ground, then at every subsequent 250 meter interval. The data are routinely reduced to

every five hundred meter increments for dissemination through various data feeds. The

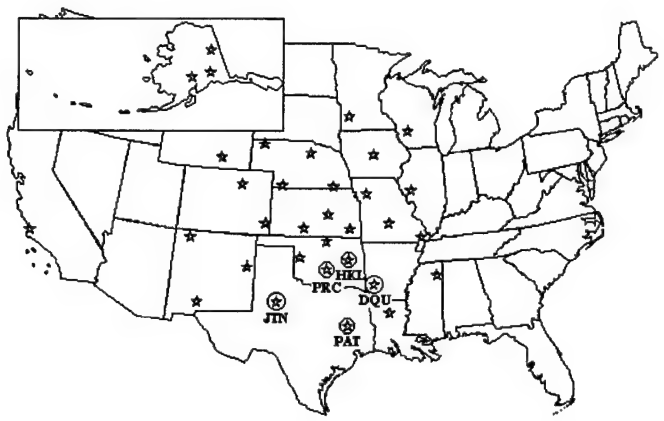


Figure 2.23: National Oceanographic and Atmospheric Administration Wind Profiler Network. Stations located within the WOP97 forecast domain are circled.

five sites located within the WOP97 inner forecast domain (indicated in Figure 2.23) include: Jayton, TX (JTN); Purcell, OK (PRC); Haskell, OK (HKL); Palestine, TX (PAT); and DeQueen, AR (DQU).

The final source of data came from pilot reports of in-flight weather conditions. Once again the main emphasis of the COMET-Tinker project is to provide AIV indicators and predictions to the operational forecasters at Tinker. The only regular source of observational data which can be used to verify these forecasts are voice-transmitted PIREPs sent from civil and military aircraft while in flight (sample shown by Figure 2.24). PIREPs reported during the valid times of the forecast runs were collected and stored for comparison with derived AIV forecasts. The use of this operational PIREP database in forecast verification has many limiting factors.

UA /OV OLM 080030/TM 1914/FL110/TP C310/SK OVC 080 CA/
TAM02/TB NEG SFC-110/IC NEG SFC-110/RM FRZLVL 078=

Figure 2.24: Sample voice transmitted pilot report (PIREP)

The main reason behind producing PIREPs is for the real-time dissemination of information to the aviation community regarding locations of hazardous in-flight weather conditions (Schwartz 1996). They are not meant as an all-inclusive observational database, and in no manner are provided to assist in rigorous research efforts by the scientific community. Shortfalls in their ability to be used for forecast verification have been identified (Shultz and Politivich 1992, Schwartz 1996, Kelsch and Wharton 1996). Examples include the formal requirements for making a PIREP which limit the information available on each report, the subjective nature in assessing aviation related weather specifics, problems regarding the spatial characteristics of the reported observations, and errors in coding the information for transmission into the National Weather Service and other databases.

The first issue identified refers to the minimal amount of information that is required to be included on all PIREPs. As listed by the governing regulations (U.S. Department of Transportation and Federal Aviation Administration 1993 – outlined in Schwartz 1996) the only required elements are a location identifier (OV), a valid time indicator (TM), a flight level indicator (FL), and the aircraft type identifier (TP). All meteorological information is treated as optional and "at the discretion" of the pilot as to whether or not it is important enough to include. The primary "additional" items that can be included are the flight level weather and visibility (WX), sky cover (SK), temperature at altitude (TA), wind direction and speed (WV), turbulence (TB), icing (IC), and any plain coded remarks (RM) to clarify or augment information already presented. To the weather community the inclusion of a flight level (the specific altitude where the aircraft is flying) temperature, wind direction, and wind speed is of importance in that it can

utilized in the analysis schemes of many of the current numerical weather prediction models. Voice reports about these items are dependent on the desire of the aircraft pilot to indicate the presence, or lack thereof, of the specific weather phenomena. The incentive for a pilot to report the presence ("Yes" reports) of icing and turbulence are primarily related to the safety of their aircraft and to that of others in the immediate vicinity. Conversely there is no real incentive for making reports of negative occurrences ("No" reports) of either phenomena. Time constraints on the pilot may limit the desire to take the necessary steps to provide a complete report, or even one at all. Therefore it is inherent in the system that there will be an underreporting of "No" icing occurrences. Conversely, "Yes" reports received will not be completely indicative of the actual distribution and frequency of icing conditions.

The second issue revolves around the ability of pilots to assess the icing and turbulence conditions that are most critical to the aviation community. The assessment of these in-flight weather conditions is subjective in nature and is primarily based on the pilot's experience, their interpretation of standard definitions and guidelines, and can be affected by the aircraft weight or type being operated. Table 2.4 gives a typical outline

Intensity	Description
Trace	Icing becomes perceptible. Rate of accumulation is only slightly greater than that of sublimation. It is not hazardous even though deicing/anti-icing equipment is not utilized, unless encountered for extended periods of time – over one hour.
Light	The rate of accumulation may create problems if flight is prolonged (over one hour). Occasional use of deicing equipment removes/prevents accumulation. It does not present a problem if the deicing/anti-icing equipment is used.
Moderate	The rate of accumulation is such that even short encounters become potentially hazardous and the use of deicing/anti-icing equipment or diversion is necessary.
Severe	The rate of accumulation is such that deicing/anti-icing equipment fails to reduce or control the hazard. Immediate diversion is necessary.

Table 2.4: Description of standard icing intensity categories (AWS 1980).

of icing intensity categories based on operating a fixed wing, reciprocating engine aircraft light and trace reports of icing in the overall PIREP database. While the experienced intensity level may be more difficult to determine than the actual presence of icing, a more definitive means of assessing icing would greatly help both the aviation and the scientific community who must rely on these reports for valuable information.

The third area of concern lies in the basic nature of the reporting of in-flight conditions. Most civil and military aircraft must adhere to mandated guidelines about how and where they can fly when traveling between two given locations. Established flight routing patterns reduce the spatial coverage which can be expected within a database of voice transmitted PIREPs. In addition, the majority of aircraft flights operating at altitudes where significant numbers of icing conditions are observed are passenger airlines which have a prescribed set of limited flight routes. Figure 2.25 shows

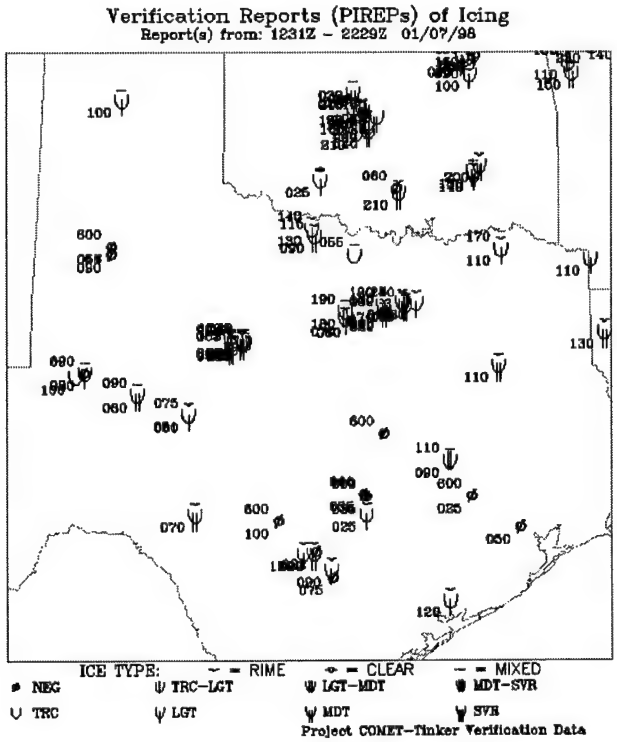


Figure 2.25: Pilot Reports (PIREPs) of icing collected between 1231 and 2239 UTC on 7 January 1998.

the icing PIREPs collected for the 1232 to 2229 UTC period of 7 January 1998. The image clearly shows that the spatial coverage of the icing PIREPs reported has marked areas with frequent reports. Other areas have little or no events reported. Indications are that the larger number of PIREPs reports occur in and around areas with airports such as the Oklahoma City and Tulsa areas in central Oklahoma, the Dallas-Ft. Worth region of north Texas and near Abilene, TX, where Dyess AFB is located. We also see that the remaining reports are typically between those locations just mentioned and other "travel" cities such as San Antonio, Houston, Midland, and Shreveport. Three concerns are raised by this examination. The first is that there may be a bias towards sampling the same icing region more than once. In instances where the forecast is correct (incorrect) in assessing the icing conditions near the heavily traveled areas, the verification result may overestimate (underestimate) the forecast's true ability. The second relates to the ability to adequately assess forecasts in all regions within the three dimensional domain. In many instances, regions where icing is forecast will not have any verifying reports. Thus the relationship between the icing forecast and the actual icing conditions may not be represented in the reports received. Finally, the compressed nature of reports may lead to conflicting reports based on icing assessments being by different pilots.

The last major concern about using PIREPs as a source for verification of forecasts of icing and other AIVs deals with administrative procedures which may lead to misrepresenting the exact location and time the event was experienced. One instance of this can be seen when a pilot reports the observed conditions after arrival at the final destination. The location and time of the observed conditions may not be reported as

precisely as if it had been made while the aircraft was in flight experiencing the event. Additionally, the locations that are typically reported on many PIREPs while in flight are the *present* position of the aircraft. Thus the reported location may or may not match the location of any of the reported "additional" data included in the PIREP. There is also the chance that, as the PIREPs get relayed from person to person, errors may be introduced. It is common for the pilot to pass a report over the PMSV to the duty forecaster, who then has the duty observer enter the information into the computers for transmission. All of these steps may introduce temporal and spatial discrepancies which will not be evident from an examination of the resulting raw PIREPs.

It is clear that there exists a need by research agencies to automate the receipt and decoding of PIREPs for use in the verification process. If the large volume of reports has to be looked at by hand, especially when dealing with those covering the entire United States, the data set would be completely unmanageable. Issues related to the automated decoding of PIREPs will be discussed later when treating the actual processing of the PIREP verification.

Chapter 3: Verification / Data Processing Methodology

Section 3.1 Forecast Verification Review

When the numerical weather forecast is completed, there is the need to provide information to a set of end users who make decisions based on what is forecast. Initial statements by users that the forecasts they received were a "good forecast" or a "bad forecast" can mean a variety of things in such a subjective context. The determination of what constitutes a good, quality forecast needs to quantified so that useful information about a specific forecasting method is obtained and comparisons between different methods can be made. Wilks (1995) identified six types of scalar measures which attempt to assess the quality of a given forecast. The six measures are: accuracy, bias, reliability, resolution, discrimination, and sharpness. Through the evaluation of individual forecasts of a given event (value) and the corresponding observed event (value), most verification methods are designed to assess one or more of these measures.

Accuracy is the measure of the average relationship between a set of specific events being forecast and the actual events that occur. It is considered the most general measure of forecast quality. The individual relationships between a set of forecast and observation pairs are combined into a single number to describe the overall forecast

quality. *Bias* is the measure of the difference between the average value of a forecast and the average value of the observed event. Forecasts which routinely predict values that are too low (or high) show a bias, even though the errors present in each individual forecasts made may different. *Reliability*, or calibration, deals with the correspondence between a specific forecast value and the average observed value which occurred. It is meant to sort the forecast/event pairs into groups based on the event that was forecast. For example, for temperature forecasts in the range of 0 to 10 °C, was the average observed value in the 0 to 10 °C range as well.

Resolution refers to the degree to which the forecasts sort the observed value into groups that are different from each other. Related to reliability, resolution compares the conditional averages of the observation for various forecast values (i.e. average observed value for each forecast that could be predicted). An example would be that if the average observed temperature is nearly the same for all forecasts, independent of whether a forecast of 10 °C or 20 °C is made, then the forecast exhibits no resolution. *Discrimination* is the converse of resolution. It gives the relationship between the average forecast value given a specific observed value. It measures the ability of the forecast to predict different outcomes when different observed values occur. If when snow and rain are observed, the forecasts are nearly always rain, then the forecast is not able to discriminate between snow and rain events.

Finally *sharpness*, or refinement, is a description of the forecasts alone, without concern for the set of observed values. Forecasts which always make a prediction which is equal to the long term climatological observed value exhibit no sharpness while those which vary widely are said to be sharp.

As stated earlier, these scalar measures each attempt to isolate some quality characteristic of the set of forecast/observed events. Use of a single measure will not give a reliable indication that a particular forecast is "good". One can produce sharp forecasts by continuously making predictions which are significantly different than the climatological mean. However, sharp forecast which are not reasonably accurate cannot be considered any good. It is therefore incumbent on the evaluator to utilize a combination of these measures to adequately assess a forecast's quality.

While these scalar measures provide a beginning insight into the ability of a forecaster, or forecast system, to provide useful information regarding a specific predicted event, ultimately we wish to show that some improvement has been made relative to some reference, or control forecast. Typical choices for the control forecasts are persistence (forecasts made by using the observed conditions at some previous time), the climatological average, or using a random prediction based on the climatological frequencies of various observed values.

The relative relationship between the reference forecasts and those we are examining is expressed in terms of a forecast skill score (SS_{ref}; Wilks 1995) , which gives a measure of the amount improvement versus the reference forecasts. The generic form of the skill score, against a specified reference forecast is represented by

$$SS_{ref} = \frac{A - A_{ref}}{A_{perf} - A_{ref}} \tag{3.1}$$

where A_{perf} is the value of the accuracy measure used if the forecasts were perfect and A_{ref} is the accuracy of the reference forecasts. It is straightforward to see that if the forecast accuracy being evaluated is equal to the reference forecast skill the score will be

zero ($SS_{ref} = 0$). Positive scores indicate improvement on the reference forecast, while a negative score means that the evaluated forecast is inferior to the reference.

The particular set of verification measures used as well as the specific methodology and numerical calculations required, are dependent on the type of forecast being made. In other words the verification measures used for forecast of variables that are continuous in nature (e.g. temperature) are different from forecasts of discrete predictands (e.g. precipitation type).

3.1.1 Verification of Continuous Predictands

The first type of forecast verification technique to be discussed is that made on quantities that are *continuous* in nature. This means that a forecast or observation can theoretically take on any value within an unlimited, or relatively large, set of possible outcomes. Examples of this are common meteorological variables used in numerical weather prediction such as temperature, wind speed, and pressure. For continuous predictands the common methodology is to use individual forecast/observation *pairs* to determine the scalar performance and skill measures. The pairs can consist of either observations/forecasts at a specific point ("go to where the data are") or through comparison of three-dimensional gridded forecasts and objectively analyzed observations.

The two common to measures of the accuracy of continuous forecasts are the *mean absolute error* (MAE) and the *mean squared error* (MSE). Mean absolute error is

the average difference (forecast error) between the forecast and observed values and is given by

$$\mathrm{MAE} = \frac{1}{n}\sum_{i=1}^{n}|f_i - o_i| \tag{3.2}$$

where f_i and o_i represent the i^{th} of n pairs of forecasts and observations. For gridded data the *n* pairs represent the corresponding forecasts and observations values for *n* grid points at a unique time, or a specific grid point for *n* separate times. The MAE will be minimized (approach zero) when the forecasts and observations are in complete agreement. In the case where MAE=0, we call the forecast *perfect*. MAE is considered the average forecast error of for the given verification data set.

The second accuracy measure is mean squared error which is given by

$$\mathrm{MSE} = \frac{1}{n}\sum_{i=1}^{n}(|f_i - o_i|)^2 \tag{3.3}$$

The MSE is the average squared difference (error) between the forecast/observation pairs. A variation of the MSE is the root mean squared error (RMSE) which is given as

$$\mathrm{RMSE} = \sqrt{\mathrm{MSE}} \tag{3.4}$$

The RMSE is useful in that it retains the units of the original meteorological quantity. In the computation of MSE and RMSE, we are squaring the forecast error, thus they are more sensitive to large errors than MAE. RMSE and MSE are limited to positive numbers and range from zero for a perfect forecast, increasing as the overall accuracy of the forecast decreases.

For vector quantities such as winds, vector versions of MAE, MSE, and RMSE (mean vector error -- MVE, mean squared vector error -- MSVE, and root mean squared vector error -- RMSVE respectively, as well as vector versions expressed as a percentage

of the observed wind; mean relative vector error -- MRVE, mean squared relative vector error -- MSRVE, and root mean squared relative vector error -- RMSRVE) can be used to provide additional information versus treating the wind speed and direction separately. The RMSVE and RMSRVE are given by

$$\mathrm{RMSVE} = \frac{1}{n}\sqrt{\sum_{i=1}^{n}\left[(\mathrm{fu}_i - \mathrm{ou}_i)^2 + (\mathrm{fv}_i - \mathrm{ov}_i)^2\right]} \tag{3.5}$$

$$\mathrm{RMSRVE} = \frac{1}{n}\frac{\sqrt{\sum_{i=1}^{n}\left[(\mathrm{fu}_i - \mathrm{ou}_i)^2 + (\mathrm{fv}_i - \mathrm{ov}_i)^2\right]}}{\sqrt{\left[(\mathrm{ou}_i)^2 + (\mathrm{ov}_i)^2\right]}} \tag{3.6}$$

where fu_i and ou_i are the forecast and observed u wind components respectively, and fv_i and ov_i are the forecast and observed v wind components.

The other scalar measure commonly used in continuous predictand verification is a bias calculation. The bias (B) is determined from

$$\mathrm{B} = \frac{1}{n}\sum_{i=1}^{n}(\mathrm{f}_i - \mathrm{o}_i) \tag{3.7}$$

where f_i and o_i are the same as described for MAE. Bias gives the average deviation of the forecast from the average observed value. In other words a temperature forecast with a "cold" bias has a negative value, while a positive value indicates a "warm" bias.

For the purpose of this work, the accuracy and bias measures were used to assess the ability of the ARPS model to predict the three-dimensional structure of the atmosphere. The main emphasis was to determine if verification of various algorithm-based icing forecasts versus PIREP-based icing observations was worthwhile. In order to expect accuracy in the icing forecasts produced, the quality of the model-derived

atmospheric state must be good. Other forecast quality measures for continuous predictands can be used to do a more complete statistical verification.

As mentioned earlier, skill scores can give a measure of the improvement of a forecast system above a reference forecast. An accuracy measure, commonly the MSE, is used as the basis for skill score determination for continuous predictands. By combining equations 3.1 and 3.3 we can get a skill score

$$SS_{ref} = \frac{MSE - MSE_{ref}}{MSE_{perf} - MSE_{ref}} = 1 - \frac{MSE}{MSE_{ref}} \tag{3.8}$$

Another skill score of historical significance is the S1 score (Wilks 1995) which was designed to measure the accuracy of gradients of pressure or geopotential height against that which can be computed from the relationship to the local wind field. *Anomaly correlation* (AC) (or pattern correlation) has also been employed on data fields to detect similarities in the patterns of departure from the climatological "norm". Instead of using the actual forecast/observed pairs, the data is converted to "anomalies" by subtracting the climatological average at each individual grid point. The AC over a particular grid point is given by

$$AC = \frac{\sum_{i=1}^{M}[(f_i - C_i)(o_i - C_i)]}{\left[\sum_{i=1}^{M}(f_i - C_i)^2 \sum_{i=1}^{M}(o_i - C_i)^2\right]^{1/2}} \tag{3.9}$$

where C_i is the climatological average value observed at the grid point given by

$$C_i = \frac{1}{n}\sum_{k=1}^{n} o_i(k) \tag{3.10}$$

The AC ranges from –1.0 (forecast and observed field are completely out of phase) to +1.0 (perfect agreement between the forecasts and observed anomalies). Additional

anomaly techniques can be used to extract position (or phase) errors when observed and forecast events are similar, but not in exactly the same location and orientation. *Principal component analysis* (PCA), also called empirical orthogonal function analysis, can be used to attempt to reduce a complex and large set of variables to a more simplified one, with significantly fewer variables. After this reduction it is hoped that much of the original variability contained in the original data set is retained, while drastically reducing the quantity of data that needs to be examined. Individual principal components can isolate phenomena from diurnal variations to predominant phases in the patterns. Normally, in the atmospheric sciences, PCA is conducted on anomaly data. Lastly, *scale separation* can be used to isolate errors on specific temporal and spatial scales, and identify those scales where the predominant forecast errors are occurring.

3.1.2 Discrete Forecasts and Contingency Tables

A second type of forecast verification technique exists when we begin to examine forecast information that is presented as a set of discrete predictands. This means that a forecast or observation can take on one and only one value within a set of possible outcomes. Unlike forecasts of temperature, dew point, wind speed, and other continuous quantities, many of the forecast elements of interest in this project were broken down into such categorical elements.

As stated earlier, the desire to eliminate the need for the duty forecaster to sift through additional "raw" mesoscale model forecast information, led to the presentation of forecast information as categorical Aviation Impact Variables (AIVs). Raw model data

were re-expressed in terms of desired forecast elements such as flight level icing, turbulence, and surface precipitation type. Certain products were also broken down further into specified intensity ranges such as trace, light, and moderate. The categories used matched the needs of the duty forecaster when preparing the Terminal Aerodrome Forecasts, Weather Watches, Warnings, and other tailored products.

To address the evaluation and verification of categorical forecast elements, the standard convention is to use a contingency table (Wilks 1995), illustrated in Figure 3.1. All categorical forecasts made by the COMET-Tinker project for this work are

		Observed (O)		
		Yes	**No**	
Forecast (F)	**Yes**	A	B	TF_y
	No	C	D	TF_n
		TO_y	TO_n	TC

Figure 3.1: Sample 2x2 contingency Table for binary forecasts. O represents observed events, F stands for forecast events, and T stands for the total number of events of a specified category. TC is the total number of forecast/event pairs. A, B, C, and D are correctly forecast "Yes" events, incorrectly forecast "No" events, incorrectly forecast "Yes" events, correctly forecast "No" events respectively.

"YES"/"NO" forecasts of the desired element. Thus the set of categorical states are represented by either a "YES" event, where the element is assessed to "exist", and "NO" events, where the element "does not exist". The term "event" in this context describes any individual observational assessment of the desired element. Thus, an observed instance of the existence of icing at a given flight level is considered a single "observed event", in this case "YES". Conversely, an observation of no icing at a point is a single "NO" event. A forecast/event pair therefore represents the combination of an observed and a forecast assessment at the same point in space, at the same time. Figure 3.1 shows

a sample 2x2 contingency table for a dichotomous forecast element, one where only two possible categorical states exist.

The main components of the contingency table are entries A through D, which are combined to assess accuracy measures in a manner similar to those for continuous forecast scalar measures discussed previously. A is the number of times the observation was correctly forecast to occur (i.e. a "YES"/"YES" forecast/event pair); B counts the times when the event was not observed but was forecast ("YES"/ "NO"); C is the number of times the event was observed but wasn't forecast ("NO"/"YES"); and D sums the times when the event was neither observed or forecast ("NO"/"NO").

The marginal sums represented by TOy (A + C), TFy (A + B), TOn (B + D), and TFn (C + D) are the total observed "YES" events, total "YES" forecasts, total observed "NO" events, and the total "NO" forecasts respectively. Lastly, TC is the total number of event pairs in the sample (TC = A + B + C + D).

A perfect set of forecasts would be indicated by C = B = 0. Since we expect some varying degree of imperfection, combinations of the main table elements and the marginal sums can been utilized to determine the overall accuracy of the forecasts.

Wilks (1995) identified the Hit Rate as the basic accuracy measure for categorical forecasts.

$$\text{Hit Rate} = \text{HR} = \frac{\text{Correct Forecasts}}{\text{Total Events}} = \frac{A + D}{TC} \tag{3.11}$$

This measure gives equal credit for correct forecasts of the "YES" event, as it does for correct forecast "NO" events, and thus satisfies the principal of equivalence of events.

In many forecasting instances, the non-occurrence "NO" event is of relatively lesser importance than the "YES" events. In these cases, an alternative to the HR is given by the Threat Score (TS), also known as the Critical Success Index (CSI; Schaefer 1990):

$$\mathrm{CSI} = \mathrm{TS} = \frac{\text{Correct "YES" Forecasts}}{\text{Total "YES" Forecasts } \textit{and / or} \text{ "YES" Events}} = \frac{\mathrm{A}}{\mathrm{A}+\mathrm{B}+\mathrm{C}} \tag{3.12}$$

We see that the threat score does not take the correct forecasts of "NO" events, D, into account. A threat score of one (1) is the best (where all forecasts fall in A), while zero (0) is the worst (where no forecasts fall in A).

The Equitable Threat Score (ETS) (Schaefer 1990) was proposed to remove correct forecasts that could be thought of as resulting from chance, or a climatology, forecast from the calculation of the TS accuracy measure. With the number of random hits defined by

$$\begin{aligned} \text{Random Hits} &\equiv \frac{(\text{Total "Yes" Forecasts})(\text{Total "YES" events})}{\text{Total Events}} \\ &= \left[(\mathrm{A}+\mathrm{B})(\mathrm{A}+\mathrm{C}) \Big/ \mathrm{TC} \right] = RH \end{aligned} \tag{3.13}$$

the equitable threat score can be expressed as

$$\begin{aligned} \mathrm{ETS} &= \frac{\text{Correct "YES" forecasts above chance}}{(\mathrm{TF_y} \text{ and/or} \mathrm{TO_y}) - (\text{Correct "YES" forecasts above chance})} \\ &= \frac{\mathrm{A} - \mathrm{RH}}{\mathrm{A}+\mathrm{B}+\mathrm{C}-\mathrm{RH}} \end{aligned} \tag{3.14}$$

This accuracy measure is a better representation of the actual ability of the forecast to provide more useful guidance than persistence, climatology, or chance. One benefit of this formulation is that "YES" forecasts of events that have a low climatological probability of occurring are not discouraged. This is made possible by the fact that as

the probability of an event decreases, the possible gain to the ETS by a correct forecast increases.

The last two accuracy measures that are typically treated for contingency tables are the Probability of Detection (POD), the proportion of times that a "YES" event was correctly forecast, and the False Alarm Rate (FAR), the proportion of "YES" forecasts that do not verify. These two measures are given by Eq. 3.15 and 3.16 respectively.

$$\mathrm{POD} = \frac{\text{Correct "YES" Forecasts}}{\mathrm{TO_y}} = \frac{\mathrm{A}}{\mathrm{A+C}} \tag{3.15}$$

$$\mathrm{FAR} = \frac{\text{Incorrect "YES" Forecasts}}{\mathrm{TF_y}} = \frac{\mathrm{B}}{\mathrm{A+B}} \tag{3.16}$$

Both range from 0 to 1, with desirable POD (FAR) at the higher (lower) end respectively. Additional information from the contingency table has been enumerated upon by Doswell et al. (1990), who outlined ways to combine the components of the 2x2 contingency table to get a complete summary of the data set. The one which will be used later in the discussion of the verification of icing forecasts using pilot reports treats the case for non-occurrences of events separately from those of occurrences. In this case the probability of "NO" detection ($\mathrm{POD_{no}}$), identified by Doswell et al. (1990) as the probability of a null event, can be treated in the same manner as the normal POD and expressed as

$$\mathrm{POD_{no}} = \frac{\text{Correct "NO" Forecasts}}{\mathrm{TO_n}} = \frac{\mathrm{D}}{\mathrm{B+D}} \tag{3.17}$$

A final verification tool for 2x2 contingency Tables is the concept of bias,

$$\mathrm{BIAS} = \frac{\mathrm{TF_y}}{\mathrm{TO_y}} = \frac{\mathrm{A+B}}{\mathrm{A+C}} \tag{3.18}$$

While not a direct relationship between correct forecasts and incorrect forecasts bias can give an indication of the pre-determination to overforecasting (bias greater than one) or

underforecasting (bias less than one) of a given event. Unbiased forecasts result in a value of one, indicating that the event occurred the same number of times as forecast. Care must be taken when interpreting the resulting bias calculation in that a set of partially or completely wrong category forecasts can still lead to a bias of one.

As noted before, all category forecasts evaluated in this work involve the basic dichotomous, "YES"/"NO", forecast that directly relates to the 2x2 contingency table measures illustrated. Other forecast products such as precipitation type, which have more than two categories into which the forecast/event can fall are still capable of using the same accuracy measures. In these cases multiple 2x2 contingency tables are evaluated for each category. For each discrete forecast category, the resulting 2x2 matrix would be achieved by treating all other forecast/event categories as "NO" cases. In the precipitation type example, with categories of rain, snow, and freezing rain, the snow forecast/event contingency Table would lump cases with rain and freezing rain as "NO" forecasts or events, with only snow cases remaining as "YES". In this manner an accuracy determination can be made for each set of collapsed contingency tables.

For square contingency tables, 2x2 for our forecasts, various skill measures can be determined to express the relative levels of improvement. The most common is the Heidke skill score (HSS; Heidke 1926) which uses the hit rate that would be achieved by random forecasts as the reference forecast accuracy measure. This hit rate is subject to the constraint that the marginal sums of the contingency that would characterize the random forecasts are the same as that of verification sample. This approach implies that there is statistical independence between the forecasts and events, or that the

determination of a forecast predictand is not dependent on what the observed event's value is, and vice versa. For our 2x2 contingency tables the HSS is given by

$$HSS = \frac{2(AD - BC)}{[(A+C)(C+D)+(A+B)(B+D)]} \tag{3.19}$$

A score of zero represents no improvement on the reference forecasts while positive (negative) values show an improvement (inferiority) to the reference, with limits of ± 1. A second skill score used frequently is the true skill statistic (TSS; Wiles 1995) or Kuipers skill score (KSS) which can be expressed by

$$KSS = \frac{(AD - BC)}{[(A+C)+(B+D)]} \tag{3.20}$$

Again the hit rate of random forecasts is used as the reference, with the difference being that the constraint is for unbiased forecasts. One benefit of this skill formulation is that either purely random forecasts or constant forecasts (always forecasting one category) will result in a KSS of zero. A second is that, like the ETS, correct forecasts of events with low probability of occurrence make larger contributions to the overall skill.

3.1.3 Adjustments for PIREP Based Verification

In Chapter 2 various concerns about using PIREPs as the source of verification data were discussed. Of critical importance in the formulation and use of the preceding scalar accuracy measures and skill scores for contingency tables is that the verification data being used obey specified assumptions. The primary assumption is that the reporting of "Yes" and "No" events be consistent. This means that both types of events are reported with the same diligence. It was shown that for PIREPs this is not true.

Recent studies by Brown et al. (1997) and Carriere et al. (1997) have attempted to assess various icing algorithms using PIREPs as their source of verification data. Accuracy measures that rely on the combination of "YES" and "NO" events into one resulting value cannot be treated in the regular manner. For example, the false alarm rate relies on the fact that both event types are equally likely to occur. From the discussion of the PIREP database we see that this is not expected to be true. There is likely to be bias toward reporting areas where icing conditions exist. In addition the limited spatial coverage of PIREPs leads to the inability to effectively verify AIV forecasts throughout the entire three-dimensional forecast domain. Thus the FAR rates resulting from a standard contingency table approach will likely be spurious and not indicative of the true measure. Alternative means of assessing verification using PIREPs as the observational data source must be introduced.

The probability of detection discussed earlier can be better thought of in this manner as the *percentage of detection* of icing (POD_{yes}). In other words, the proportion of "Yes" PIREPs that were located at a point where the forecast was also "Yes". Conversely, we can treat the probability of "No" detection as the *percentage of "No" detection* (POD_{no}). It is possible in this case to create multiple indicators of POD_{yes} and POD_{no} by treating different subsets of the event spectrum individually. In the example of icing verification the ability of the algorithm to detect area with reports of moderate or greater icing may be of more significance than simply detecting all icing events.

Another set of accuracy measures we can consider are based only on the forecasts. We can always assure a perfect POD_{yes} if the forecast always predicts "Yes". This does not provide any usable information since we would like the algorithm to help the

forecasters isolate regions where the event is "particularly likely" to occur. In fact, overforecasting can be detrimental to the entire forecast process by causing the forecaster to lose faith in the forecast model altogether. A measure of the total region of the model domain with a "Yes" forecast could be used to see if overforecasting is occurring. To measure this, two quantities are introduced, the Impacted Area (IA) and Impacted Volume (IV) of a forecast (Carriere 1997; Brown 1997).

The Impacted Area is defined as the area projected along a horizontal surface where at least one forecast level in the vertical has a "Yes" forecast. Using a two-layer model (i.e. two vertical levels k=1 and k=2) as an illustration, with forecast areas as indicated in the top two panels of Figure 3.4, the resulting IA would simply given by the surface area covered by *both* forecast regions.

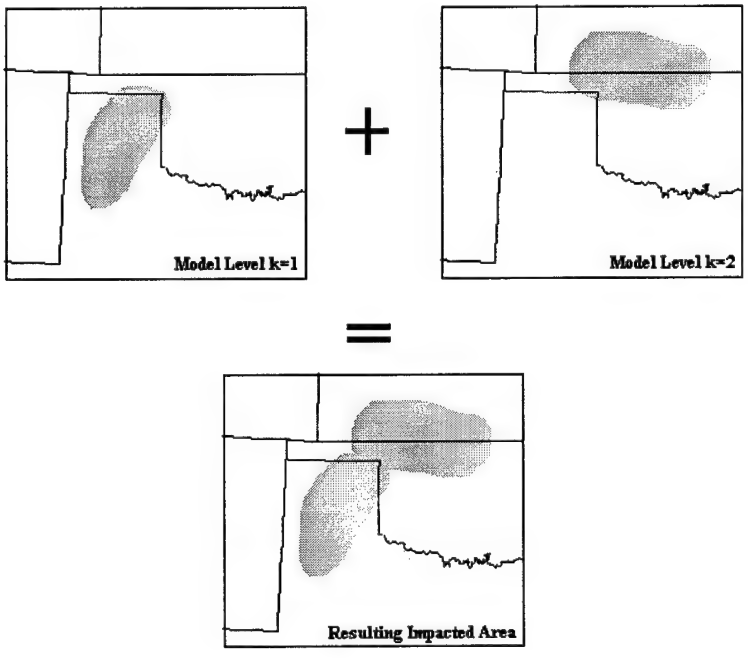


Figure 3.4: Illustration of Impacted Area (IA) assessment. Top two panels show the individual level forecasts (shaded area indicates a "Yes" forecast) in a two layer model. The bottom panel shows the resulting IA.

In a similar manner the IV is defined as the sum of the three-dimensional volume surrounding each grid point with a "Yes" forecast. In an absolute sense, differing values of either IA or IV are not indicative of better or worse performance. From a forecaster standpoint, smaller values may be more desirable so as to remove a greater portion of the forecast domain from consideration. However, since PIREPs do not adequately sample the entire domain, larger forecast regions may actually be better assessments of the true distribution of "Yes" conditions at any given instance. We can expect, however, that as the amount of area and volume where the event is forecast to occur increases, the likelihood of overforecasting increases.

The percentage of detection measures and the impacted area/volume calculations can be combined to give a third measure of the efficiency of the forecast. We can define two ratios, the area efficiency (AE) as

$$AE = \frac{POD}{IA} = POD \text{ per unit area} \tag{3.21}$$

and volume efficiency (VE), as

$$VE = \frac{POD}{IV} = POD \text{ per unit volume} \tag{3.22}$$

as additional indications of the relative abilities of individual forecasts to accurately predict locations of observed "Yes" events. It is easy to see that the AE and VE can be improved with either an increase in the detection rates and/or a decrease in the impacted area/volume. Thus, in a simplified manner, higher efficiencies indicate better algorithm performance relative to those with lower values.

Again, the AE and VE calculations cannot be treated as absolute measures of algorithm performance. Primarily, the size of the forecast domain alters the absolute

magnitude of the calculated measures. In addition, varying efficiency calculations can arise from different means. Algorithms with low POD capabilities can still look extremely efficient if the region of icing forecast is exceptionally small. Whether the forecast provides useful information and can provide additional skill to the forecast process is a subjective assessment. It is the interpretation of the entire set of identified measures together which is needed to make a general ability determination.

Section 3.2 Point Observation Comparisons: Surface / Wind Profiler

The first type of verification performed involved the comparison of forecast output to data collected from the various surface-based observations sources mentioned earlier. Mewes (1997) developed an initial verification suite (referred to as "original code") for use in evaluating the ARPS forecast model. It is designed to aid in both real-time verification and case study analysis. The components of the code used in this section interpolates the raw model variables to the exact location in three-dimensional space where the observation was taken. The model variables can then be converted into the quantities that are directly measured by the observing equipment.

Profiler-based wind measurements were collected for the forecast period for the five stations listed previously. The model-derived u and v wind components were interpolated to the vertical column directly above the identified latitude and longitude (lat/lon) of each profiler location. A one-dimensional Barnes analysis (Barnes 1973) was then performed to interpolate from the vertical levels used within the model to the standard profiler observing heights above the ground level (500 m increments). The

resulting wind components were then simply converted into a forecast wind direction and speed at each level.

In order to prepare the observed data for the verification code, a quality control check was performed to ensure erroneous data were eliminated. If the observed data were not at the proper height levels a Barnes analysis similar to that used for model data was performed. Then a basic vertical consistency check (Mewes 1997) with adjacent levels was performed to remove data which may be in error. Computations of the bias and RMSE errors for both the wind speed and direction were made using each corresponding forecast/observation pair with the results averaged over the entire set of forecasts, stratified for each specific valid hour, and indicated reporting levels.

The verification against surface observations is a bit more complex. One reason is that the model variables and those that are actually observed are different. Another is that the vertical resolution in the low levels of the model does not predict these variables at the *anemometer level* or at the *surface*, but at some point about that. Large vertical gradients of model derived quantities may lead to large errors if interpolation is used to obtain model values near the surface. For this purpose the original verification code incorporated the Monin-Obukov surface layer similarity theory (as described in Mewes 1997), which uses stability dependent profiles of wind shear and temperature gradients to reduce the temperature and wind speed forecasts within the model "surface layer" to any particular height of interest. For a surface comparison, the heights are the World Meteorological Organization standard 2 meter (temperature) and 10 meter (winds) heights. The other surface observation variables which are not covered by the similarity theory (e.g. dew point and pressure) were reduced to the surface by linear interpolation or

other adjustment (i.e. hydrostatically for pressure) between the two lowest model levels. In the ARPS configuration used here, the lowest level (k=1) is ten meters below the ground while the second level (k=2) is the same distance above. The terrain in the ARPS model is interpolated to each desired horizontal grid location from a 30 second by 30 second data base developed by the United States Defense Mapping Agency which may introduce small errors in model terrain. In some instances, the model derived terrain values have errors on the order of 10 meters. While not a large difference, preliminary evaluation of the surface pressure verification from WOP97 showed that this did affect the overall results. To correct this, the verification code was modified to compare the model derived terrain with actual station elevations and adjust the forecast station pressure by using a hydrostatic correction. The final result was a set of model-derived "estimates" of temperature, equivalent potential temperature, dew point, wind speed and direction, and station pressure.

Little post processing of the actual surface data was required. The routine surface observations were already in the desired units and format. The Mesonet observations of relative humidity were converted to dew point temperature. As with the wind profilers, bias and RMSE calculations were made from each corresponding forecast/observation pair. The results were then averaged over the entire set of forecasts for the various surface locations, and for unique subsets of the forecast period.

Section 3.3 RUC Field Comparisons

The original field verification code developed by Mewes (1997) was designed to

operate on a single forecast run. For this reason, it was designed to interpolate the gridded verification data to the ARPS model grid and perform the verification on the model levels. While this can be done for a set of daily forecasts, a more conventional constant pressure level comparison would allow easier interpretation by forecasting personnel. Using the RUC analysis data already identified, comparisons with the ARPS forecast were made at standard pressure surfaces (850, 700, 500, and 300 mb) as well as at the surface.

For the ARPS forecasts, a linear interpolation of model variables using a logarithmic pressure coordinate was performed. Conversion of the model variables on each pressure surface was made to obtain geopotential height, temperature, u and v wind components, dew point, and relative humidity. The variables were chosen since they were available in both the ARPS forecast and RUC analysis data, and are routinely used in forecasting operations. We wish to evaluate relative humidity since it is used in the determination of the icing forecasts. The RUC analysis data were converted to the ARPS forecast domain and grid spacing using EXT2ARPS, the primary program developed by CAPS to perform boundary condition generation and other conversions of outside sources of gridded data. The resulting files were then run through the same interpolation routine as the forecast output to generate observed variable fields on the constant pressure surfaces.

Section 3.4 Icing Forecast Comparison

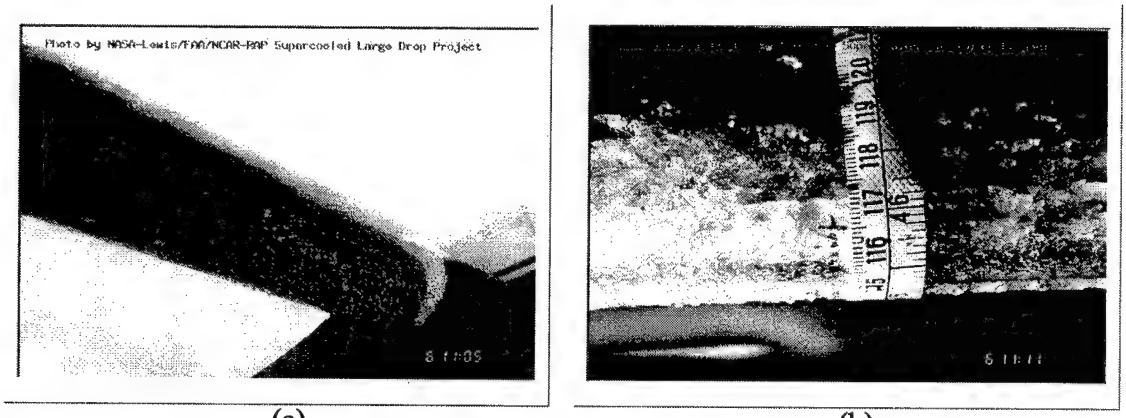
Aircraft icing is one of the primary in-flight hazards to commercial, military, and

general aviation aircraft. In recent years, special emphasis has been placed on the aviation and weather community to better evaluate icing formation mechanisms (Politovich 1989; Bernstein 1997), to determine how it affects aircraft operations, and to successfully predict regions where icing will occur (Tremblay et al. 1996; Carriere et al. 1997, Brown et al. 1997). Icing accumulation on an aircraft in flight has a wide range of impacts. Primarily, it alters the shape of the wing, affecting its ability to produce the required lift to stay in the air, and increases the overall drag of the aircraft, reducing its operating speeds and ability to maneuver effectively (Politovich 1996). Additionally, ice buildup modifies the weight and shape of the entire aircraft, leads to engine malfunctions when ice is ingested into the engine itself, and alters the operation and handling characteristics of the many aircraft control surfaces needed by the pilots. All of these can have negative impacts on the ability of the pilot to properly maintain aircraft safety in flight. Accurate forecasts of icing conditions could reduce or eliminate the problems and emergencies that are currently experienced when in-flight icing is encountered.

Two basic conditions must be met for icing to form on the structural components of an aircraft. The surface temperature of the aircraft must be at or below the freezing point (0 °C) and there must exist some supercooled liquid water (SLW) droplets (i.e. liquid water below the freezing point) in the local environment. The existence of SLW is possible since water in the free-atmosphere does not always freeze at 0 °C, but typically in the range from –10 to –40 °C (AWS 1980). The purity and size of the water droplets that exist are the primary factors in determining the actual freezing point.

Standard descriptions of icing intensity and types have been developed for reporting and forecasting of icing conditions. For in-flight operations there are two types

of icing that can be encountered, rime and clear icing. Rime icing is generally formed by the instantaneous freezing of water droplets on the skin of the aircraft and produces a rough, milky white, layer of ice. The photographs in Figure 3.3 show how rime icing appears on the surface of an aircraft. Air bubbles trapped in the ice as the water freezes



(a) (b)

Figure 3.3: Illustrations of rime icing on (a) a NASA operated DeHavilland DHC-6 Twin Otter research aircraft and (b) an aircraft wing sample after exposure to actual in-flight icing conditions

cause the milky appearance. The roughness and non-uniform buildup of this type of icing makes it a greater hazard to in-flight operations. The other type, clear icing, is formed when the water slowly freezes on the aircraft skin, allowing it to spread out more evenly before it freezes. Since freezing takes longer to occur, it cannot trap as much air and results in a clear sheet of ice (AWS 1980). With the icing distributed over a larger area of the aircraft it does not have as drastic of an impact to flight safety unless it is allowed to build continuously for an extended period of time. In some instances both types can exist at the same time, in which case the icing is considered "mixed".

Icing in clouds is affected by the physical state of water that makes up the cloud (such as liquid versus ice crystals), the size and spatial distribution of cloud particles, and

the existence of any precipitation. Clouds composed of ice, versus liquid water, are not likely to produce regions of icing that are hazardous to aircraft operations. On the other hand, clouds with vast amounts of liquid water near or below freezing (i.e. supercooled liquid water) are of primary concern. Thus the ability of any icing algorithm to accurately forecast icing is highly dependent on its ability to assess the presence of clouds, and the amount and state of water that makes up those clouds.

A caveat to the preceding is that some reports of icing conditions have been shown to exist in regions where the ambient temperature is above freezing or where there are no clouds present at all. One must remember that other, non-weather factors, do have a bearing on whether icing conditions can exist. Issues such as the recent flight-path history of the aircraft, the aircraft type, and differences in piloting techniques and procedures may all lead to variations in when icing conditions can develop. The easiest example to visualize is when a plane has been flying in a region of sub-freezing temperatures, then enters a cloud with liquid water at temperatures slightly above freezing. The skin temperature of the aircraft may remain cooler for extended periods due to sub-zero temperatures of the fuel stored in the wings. This may enable water droplets to freeze on contact at higher ambient air temperatures than normal. It is beyond the ability of most algorithms to catch all possible cases of icing formation, especially when associated with these non-meteorological factors.

In addition to the comparison of standard meteorological variables, WOP97 provided a unique opportunity to evaluate the ability of the ARPS model to predict winter-time weather phenomena. While in-flight aircraft icing occurs year round, we expect significantly more icing events, and thus pilot reports, to be observed during the

winter months. For this reason, icing forecasts comprised the bulk of the AIV products initially incorporated into the forecast suite for WOP97.

The individual PIREP data files were collected in the format described in Chapter 2 and archived for use in the verification. The entire database included nearly two hundred thousand reports made by either voice transmission or by automated means. The decoding of these reports was accomplished using a Perl routine written by Gregory Thompson (1995). Specific identifiable features of the database make aspects of the decoding procedures an issue when verifying PIREPs. These features can cause significant difficulty in decoding the reports and can introduce errors when extracting the weather information they contain.

The most difficult part of the overall decoding process is how to handle missing or inappropriate data. Examples of these types of problems are: the reporting of two station identifiers in the location block; not including a flight level or using non-standard abbreviations such as DURC (during climb); and using ground level (AGL) to report observed phenomena instead of with respect to mean sea-level (MSL). Thus PIREP data may be assigned to the wrong point in either time or space. In the situation where no flight level is indicated, the decoder attempts to use other indicators to assess a flight level to the report. An example of this is shown in the PIREP in Figure 3.4. which has no numeric flight level reported. The decoder first attempts to use any three digit numeric

GUP UA /OV GUP/TM 2253/FLDURC/TP BE90/IC NEG/A2988=

Figure 3.4: Sample PIREP with improperly coded flight level information.

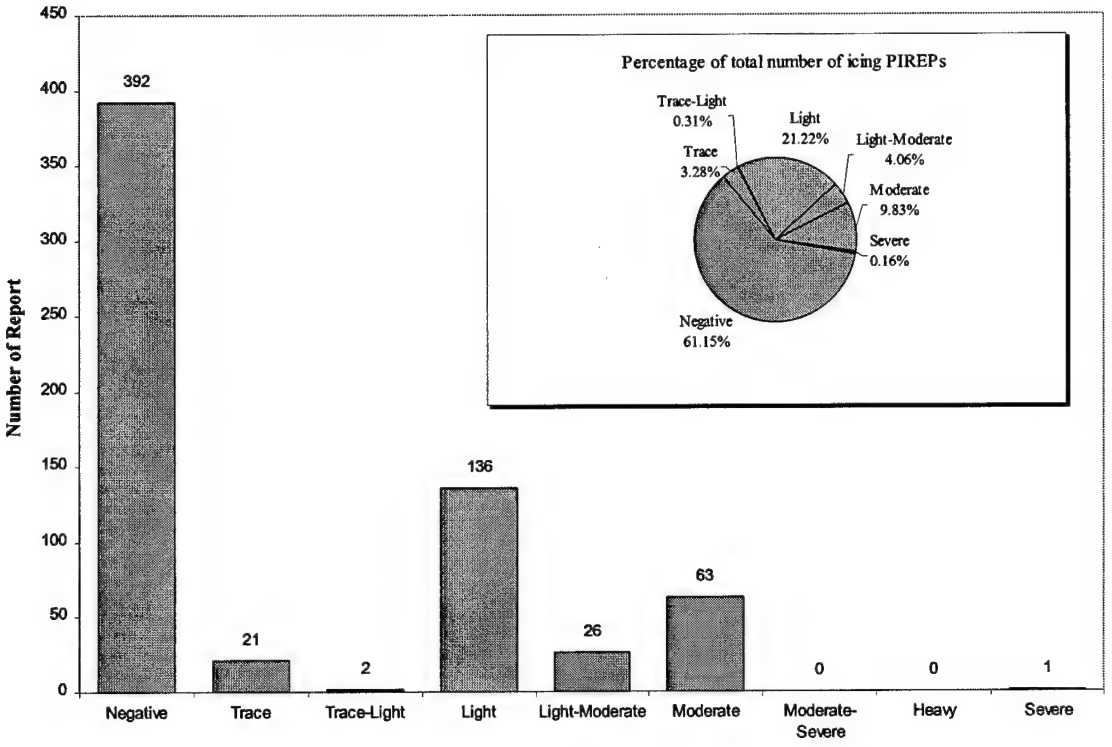
information in the turbulence, icing, or sky conditions blocks. If no levels are present, it then uses numbers in the remarks section to determine a level. The resulting flight level assesment the decoder makes in this case is to incorrectly assign the first three digits of the altimeter setting, 298, as the flight level and thus the altitude of the negative report of icing. This same procedure is followed if the time indicator is incorrectly given. A second occurs when the location identifier (OV) is coded with two locations, meaning "between" point one and point two (e.g. TIK-SAT for enroute between Tinker and San Antonio). The decoder will use the midway point between these two location as the approximate place where the PIREP is valid to assign it a single location. Lastly, all altitude assignments in the PIREP are assumed to mean "above mean sea level (MSL)". On many PIREPs made in the lower regions of the atmosphere it is a common practice to give altitudes "above ground level". The decoder has no way to recognize this and will use the level indications as if in MSL. For this reason a minimum flight level threshold (FL040 – four thousand feet) was used to remove these PIREPs from consideration in the verification. While this removed a significant portion of the database, it is believed that this practice is better than comparing reports to forecasts at the wrong height in the model.

Another factor to consider when using PIREPs for verification purposes, specifically for icing forecasts, is the fact that when icing is missing from the report it does not mean that icing was not observed. In some instances, proxy information can be used to treat the absence of icing in the report. Indications such as temperatures above freezing and the remark "Clear Above (CA)", which means there are no clouds, are possible ways to infer the absence of icing. Current decoder practice is such that reports

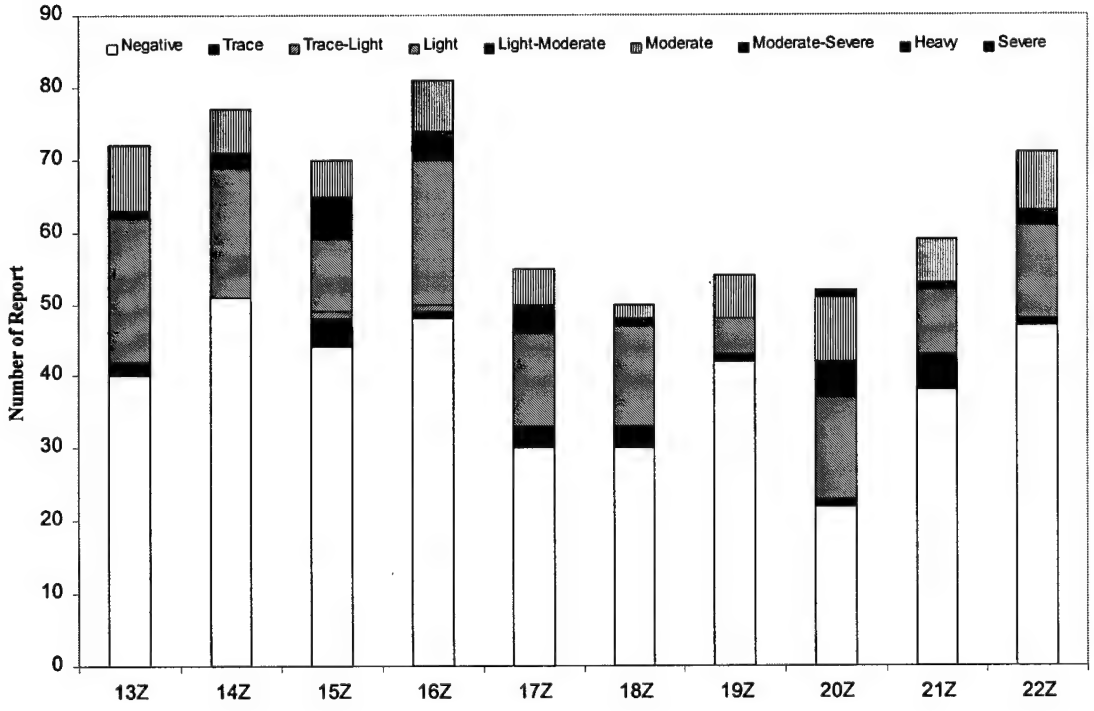
with CA indicated will be assigned a "Negative" icing report and be included in the verification database. The vertical extent will go from the flight level determined to an arbitrary domain height (currently FL600 – sixty thousand feet). The temperature restriction of above freezing is not always indicative of the absence of icing conditions (as was discussed) so these reports are not identified as "No" icing for our purposes.

During the forecast period, over 85,000 voice pilot reports were made in the United States and decoded. Of this amount, only 1734 (2%) were within the 9 km forecast domain shown in Figure 2.16 and above the cutoff flight level of four thousand feet. Of the reports in the domain, there were just 641 reports (37%) with specific indications of icing conditions. Figure 3.5(a) shows a basic breakdown of the icing reports received which indicate that they were predominantly of the "No" type (392 reports; 61%), or in the range from Light to Moderate (225 reports; 35%). In addition, Figure 3.5(b) shows good temporal coverage over each forecast hour. Each hour increment had roughly 50-80 icing reports. Figure 3.6 shows the data for each forecast day, giving the total number of icing reports (3.6a) and indications of the various number of reports at each intensity level (3.6b). We see that the icing PIREPs generally cover the full range of WOP97 forecasting days, with none having no reports at all and only a few missing "Yes" reports completely.

The final issue to examine is the assessment of how to group individual PIREPs with the applicable forecast valid time when performing the verification procedures. Since the archived forecast output is valid on every hour, PIREPs within -29 and +30 minutes of the valid hour were used in the verification of each forecast.

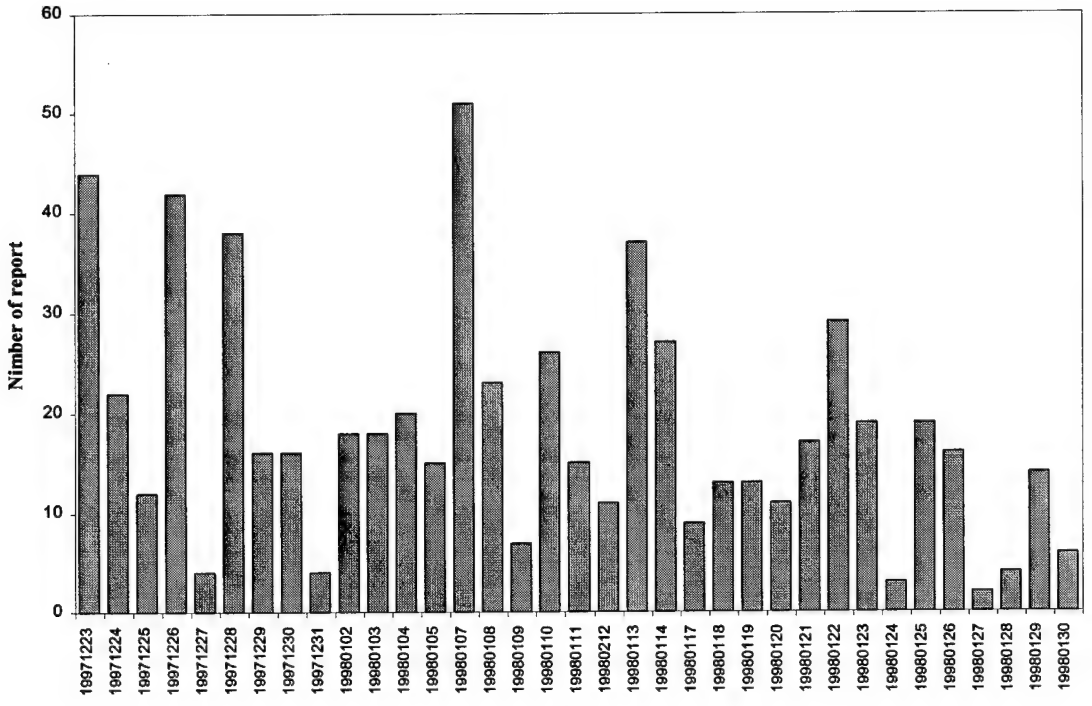


(a)

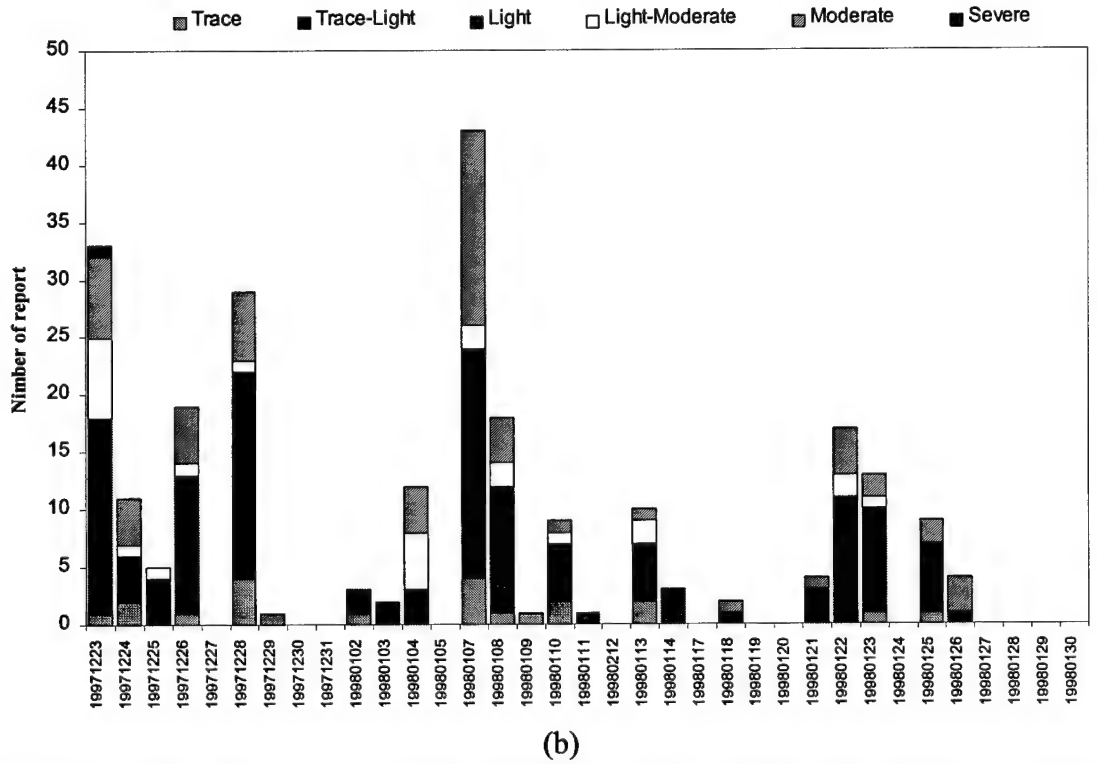


(b)

Figure 3.5: Number of Pilot Reports (PIREPs) of icing during the WOP97 grouped by (a) intensity alone (inset shows percent of all icing PIREPs reported) and (b) by the intensity and hour used for verification purposes. (Only reports meeting quality control limitations are included)



(a)



(b)

Figure 3.6: Number of Pilot Reports (PIREPs) of icing during the WOP97 grouped by date for (a) all "Yes"/"No" icing reports and (b) "Yes" only reports (intensity indicated by color shading)

3.4.1. Icing Algorithms

The ability of numerical weather prediction systems to generate a three-dimensional picture of atmospheric temperature and water content has led to the development of various computational algorithms designed to assess icing conditions from model output. For the COMET-Tinker project, six algorithms have been introduced into the ARPS forecast display software for presentation to duty forecasters. A seventh has been added for evaluation purposes but is not currently included into the forecast display system due to its similarity to other algorithms already available. The algorithm methodologies fall into two primary types. The majority attempt to assess locations of "cold clouds" in the model domain using thresholds of standard meteorological variables (i.e. temperature, dew point, etc.). The second type attempts to assess regions where supercooled liquid water exists from the microphysical parameterizations schemes within numerical models.

A number of the algorithms utilized have intrinsic icing "*type*" determinations, or are set up so that the icing forecast can be attributed to a specific cause within the algorithm methodology. Where applicable these "type" characteristics were retained on the real-time product display. This allows the forecaster the ability to make real-time modifications of the icing forecast if the observed atmospheric conditions do not match with the expected cause of the icing. For evaluation purposes and comparison with observed PIREPs, all algorithm results were reduced to a standard yes/no icing forecast.

(a) *Threshold Based Icing Algorithms*

The first set of algorithms are designed to be first guesses of locations of icing based on the idea that icing only exists in clouds, at specific temperature ranges. Most have been modified from existing sounding-based or empirically-based flow charts that have been routinely used by weather forecasters. The original intent was to apply these algorithms to observations or model output that did not included explicit microphysical parameterizations or variables. The algorithm computations are done at each three-dimensional grid point in the forecast domain to get the final forecasts of expected icing regions. One drawback to these "threshold based" algorithms is that they normally combine empirically determined thresholds with simplified theoretical assumptions about how and where icing occurs. Even though icing is not assumed to exist in sub-saturated air or cloud-free regions, many instances of relative humidity (RH) below 100% in the model will actually correspond to observed cloudy regions. This is due to the assumptions made in the model physics (e.g. sub grid scale processes parameterizations), the spatial resolution used in the model calculations, and errors inherent in the calculations themselves. Since the model predictions of needed meteorological variables are subject to processes that cannot be completely resolved by the model, RH values lower than 100% are often used to "assess" cloudy regions from the model output. This use of relative humidity (or dew point) as a criteria for cloud existence is an oversimplified methodology that may lead to overforecasting regions of icing. In addition, clouds that are made up of primarily snow and ice may still result in an positive icing forecast since the algorithms have limited means to discriminate these from liquid water clouds.

(1) *The RAOB Icing Algorithm*

The first algorithm of this type is one that was initially developed by the Air Force Weather Agency (AFWA, formerly the Air Force Global Weather Center – AFGWC) for the RAOB rawinsonde data plotting software program. For simplicity it will be referred to as the RAOB algorithm for the remainder of this discussion. The RAOB algorithm is shown in complete form in Table 3.1.

The premise behind the RAOB algorithm is that icing conditions will exist in cold clouds, which are indicated by dew point depressions ($T\text{-}T_d$) in the range, $0 < T\text{-}T_d < 4$ °C, where the ambient air temperature is below freezing. The type of icing that will exist is primarily determined by the lapse rate between the grid points directly above and

Temperature, °C	$0 \geq T > -8$				$-8 \geq T \geq -16$				$-16 \geq T \geq -22$
Dew point Depression, °C ($T\text{-}T_d$)	≤ 1		$1 < T\text{-}T_d \leq 2$		≤ 1		$1 < T\text{-}T_d \leq 3$		≤ 4
Lapse Rate °C / 1000 ft	≤ 2	> 2	≤ 2	> 2	≤ 2	> 2	≤ 2	> 2	N/A
Icing INT	LGT	MDT	TRC	LGT	MDT	MDT	LGT	LGT	LGT
Forecast TYPE	Rime	Clear	Rime	Clear	Rime	Mixed	Rime	Mixed	Rime

Table 3.1: Overview of AFWA RAOB Icing Algorithm based on temperature (T; °C), dew point depression ($T\text{-}T_d$; °C), and lapse rate (°C /1000 feet) thresholds. (TRC – Trace, LGT – Light, and MDT – Moderate)

below the point where the icing forecast is to be assessed). The algorithm does not assess any icing potential when the air temperature is below –22 °C since all water is assumed to be in an ice state and not able to facilitate icing formation. Figure 3.7 shows

a sample two-dimensional forecast display of the RAOB icing algorithm at a specified height above mean sea-level (MSL) for the domain used in the winter operational test period. Intensity is indicated by the color shading while the character overlay shows the icing type.

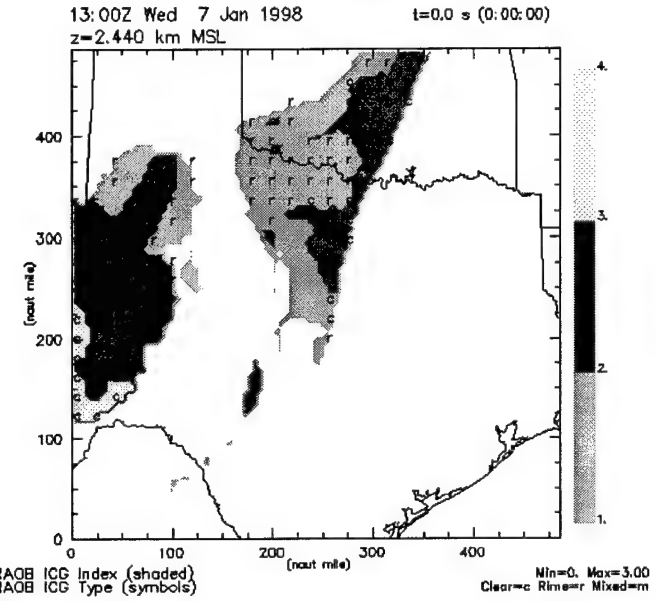


Figure 3.7: Sample AFWA RAOB Icing Algorithm forecast product. Intensity indicated by the color shading (1 – Trace, 2 – Light, 3 – Moderate), type indicated by text (c – Clear, r – Rime, and m – Mixed)

It has been shown that the icing type determined by the RAOB algorithm does not display any real skill in forecasting icing type (Cornell et al. 1995), compared to observed pilot reports. For the real-time display of products to forecasting personnel the type assessment was included to allow use in a general subjective manner. As a final point, the RAOB algorithm does not include criteria to forecast "severe" icing, which is of extreme importance to the aviation community.

(2) *The NCAR/RAP Icing Algorithm*

The second threshold based algorithm was initially developed by the National Center for Atmospheric Research / Research Applications Branch (NCAR/RAP) (Schultz and Politovich 1992; Forbes et al. 1993) and is under continuous refinement

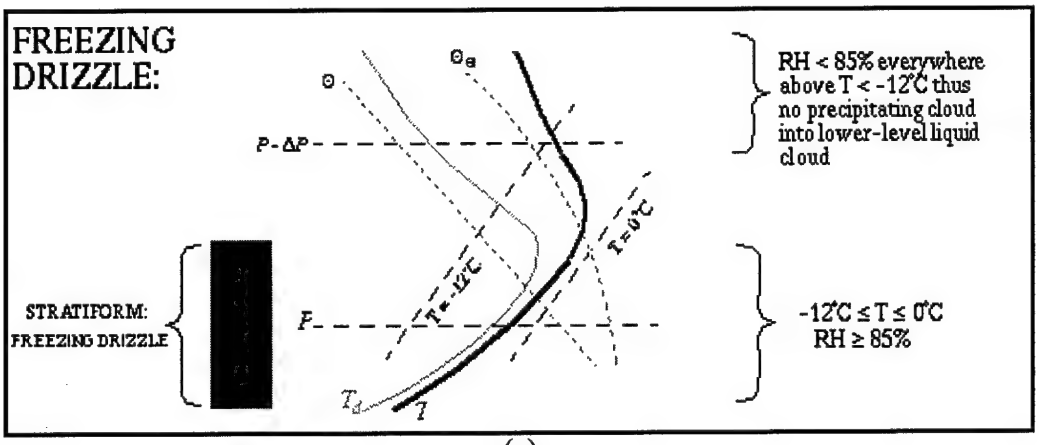
Temperature, °C	$-12 \leq T \leq 0$ °C	$T \leq 0$ °C	$-20 \leq T \leq 0$ °C	$-16 \leq T \leq 0$ °C
Relative Humidity, %	≥ 85 % (< 85 % above where $T < -12$°C)	≥ 80 % (≥ 80 % above top $T > 0$ °C level)	$82\% \geq RH \geq 63$ % (Max in conditionally unstable layer ≥ 63 %)	$82\% \geq RH \geq 63$ %
Icing ("Yes" forecast)	Stratiform (Warm Stratus)	Freezing Rain	Unstable	General (Stable)

Table 3.2: Overview of NCAR/RAP Icing Algorithm based on temperature (T; °C) and relative humidity (%) thresholds.

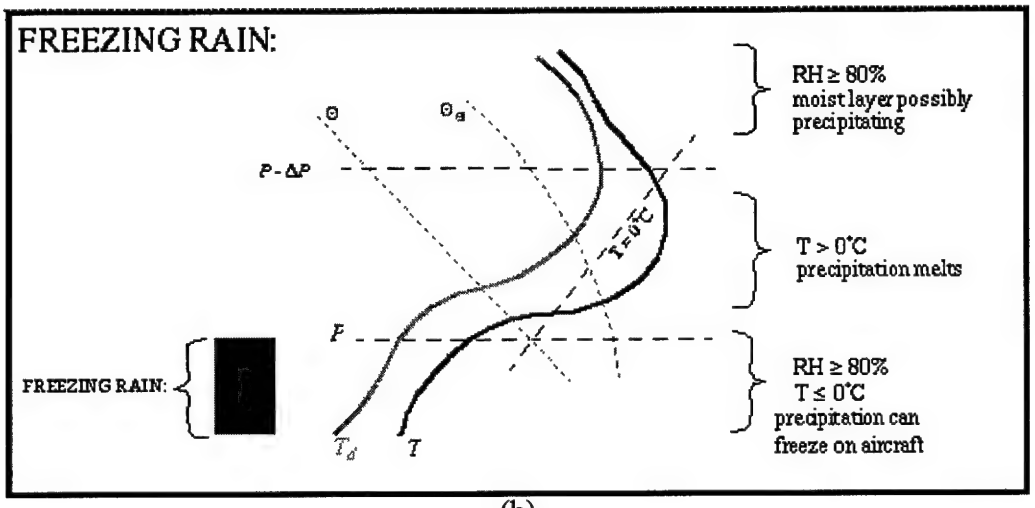
(Thompson et al. 1997). The inclusion into the COMET-Tinker forecast suite started with Version 4 of the algorithm, which allows four separate categories for icing to exist. A summary of the algorithm as used by COMET-Tinker is shown in Table 3.2.

The first scenario, the stratiform portion (illustrated in Figure 3.8(a)), attempts to identify regions of "warm" stratus clouds with ambient temperatures from 0 to –12 °C and RH ≥ 85%. No layer with relative humidity above 85% and the air temperature less than –12 °C can exist above the tested level. This dry upper layer requirement is designed to prevent higher clouds from precipitating ice into the low clouds and removing the available liquid water.

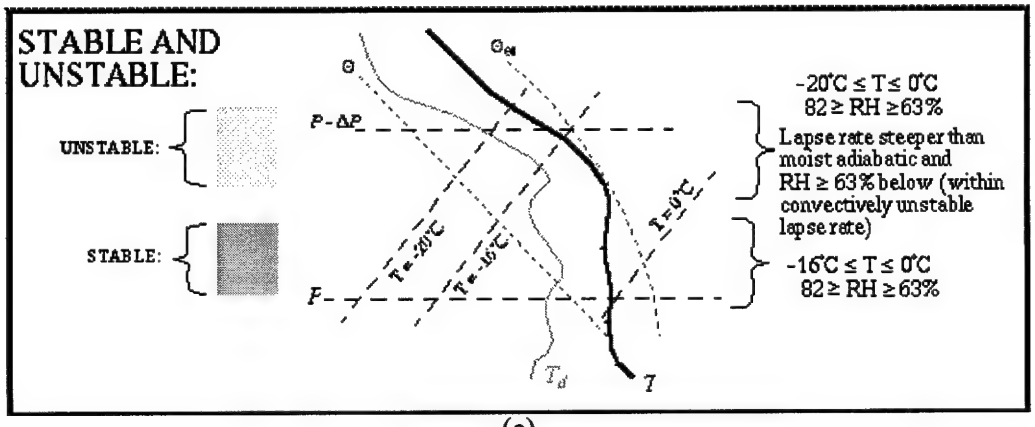
The second case, the freezing rain scenario, is intended to catch typical Midwestern and Northeastern United States freezing rain events. A layer of warm air, with T > 0 °C, is above a region of temperatures below 0° C. If there exists a precipitating cloud above these two layers (indicated by RH ≥ 80% somewhere above the point being evaluated) and a sufficient amount of water in the warm layer itself (RH > 80



(a)



(b)



(c)

Figure 3.8: Illustration (partial Skew-T diagram) of the NCAR/RAP Icing Algorithm – (a) stratiform (freezing drizzle) icing scenario, (b) freezing rain icing scenario, and (c) stable and unstable icing scenarios. The solid lines are the vertical temperature (T) and dew point (T_d) profiles as indicated. (after Thompson et al. 1996)

%) so that the precipitation melts but does not evaporate, icing may then be present in the lowest layer where the air temperature is below 0 °C . The freezing rain scenario is illustrated in Figure 3.8(b). There are two areas of concern that aren't checked with this simplified freezing rain scenario. The first is a check to ensure that the warm layer is of sufficient depth to actually allow the precipitation to melt. The second is to check that the vertical distance between the precipitating upper layer and the start of the warm is small enough so that the precipitation does not evaporate before reaching the warm nose.

The general (stable) icing category is designed to catch large-scale cold-cloud regions (clouds with air temperatures in the –16 to 0 °C range) and is similar in nature to the RAOB algorithm and the original Schultz and Politovich (1992) scheme. The relative humidity threshold varies linearly from 82% when the temperature is at or near 0°C to 63% for temperature approaching –16 °C. The stable scheme is illustrated in Figure 3.8(c).

The final portion of the NCAR/RAP algorithm, the unstable category, is designed to catch intermittent icing in areas susceptible to smaller scale convection, such as low level strato-cumulus, or convective elements within larger scale stratiform regions not matching the other icing criteria. There must exist a conditionally unstable layer in the model immediately below the grid point where the forecast is valid. If a conditionally unstable layer is present, the air temperature and relative humidity checks are accomplished. The ambient air temperature must be between 0 and –20 °C. The relative humidity threshold varies linearly, similar to the general icing case, from 82% when the temperature is at or near 0°C to 63% for temperature approaching –20 °C. The unstable scheme is also illustrated in Figure 3.8(c).

Figure 3.9 shows a sample two-dimensional forecast display of the NCAR/RAP icing algorithm at a specified height above MSL similar to Figure 3.7. The NCAR/RAP

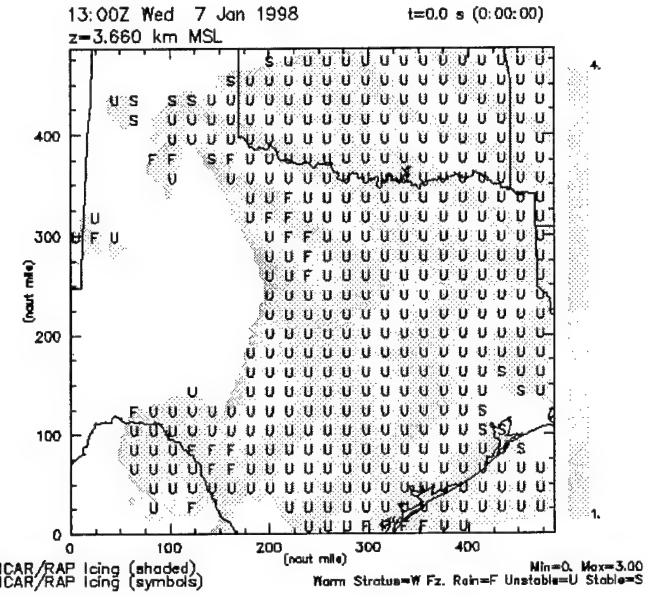


Figure 3.9: Sample NCAR/RAP Icing Algorithm forecast product. Existence of icing indicated by the color shading, type indicated by text (W – Warm Stratus, F – Freezing Rain, S – Stable, and U – Unstable)

product available to the forecaster shows the assessed icing type, indicating the scenario that caused the icing to be forecast.

(3) *The AWC Icing Algorithm*

The third threshold-based algorithm was originally developed for use by the Aviation Weather Center (AWC, formerly the National Aviation Weather Advisory Unit – NAWAU) of the National Weather Service (NWS) as guidance for AIRman's METeorological Information (AIRMET) preparation (Thompson et al. 1997). The AWC provides analysis and forecast information regarding weather conditions that will affect domestic and international aviation interests. The AIRMETs are designed to advise for

weather that may be hazardous, other than convective activity, to single engine and other light aircraft. This product is primarily geared to Visual Flight Rule (VFR) pilots who have restrictions on where they can fly.

The AWC algorithm (shown in Table 3.3) is also based on the Schultz and Politovich scheme (1992) from which the NCAR/RAP algorithm evolved. It assesses two different icing potential categories, high and low probability. The AWC algorithm is not currently in the real-time production suite generated by COMET-Tinker, since it is of similar nature to the RAOB and NCAR/RAP algorithms already encoded.

Temperature, °C	$-14 \leq T \leq -1$ °C	$-20 \leq T \leq 0$ °C	$-19 \leq T \leq 0$ °C
Relative Humidity, %	≥ 75 %	≥ 86 %	≥ 60 %
Altitude (MSL), meters	> 900 m	$\leq$ 900m	All
Icing Forecast	High Probability	High Probability	Low Probability

Table 3.3: Overview of AWC Icing Algorithm based on temperature (T; °C) and relative humidity (%) thresholds.

(4) ***The LAPS Icing Index Algorithm***

The fourth threshold-based algorithm is based on the Local Analysis and Prediction System (LAPS) Icing Severity Index and other code developed for LAPS. It is slightly different from the three algorithms just mentioned in that it assesses the presence of clouds through a computed liquid water content (LWC). Hence, it requires a model with explicit microphysics or other water content determinations.

LAPS was developed by the Forecast Systems Laboratory as a completely integrated analysis, data assimilation, and forecast system (Albers et al. 1996), with the ingest of various sources of data including surface observations; mesonet reports;

automated aircraft data (ACARS); radar; satellite; and others. LAPS incorporates all the available data into a single three-dimensional gridded database for initialization of mesoscale forecasts. Attempts have been made to retain as much of the original information as possible. For this reason, specific algorithms have been designed to assess liquid water content, three dimensional cloud cover and cloud type, precipitation type and rate, cloud drop size and density distribution, icing severity and other non-standard meteorological. Updates to the LAPS code for these non-standard features have been ported to ADAS, the data analysis system in use by ARPS.

Of primary interest to COMET-Tinker is the icing severity index, which requires three-dimensional fields of liquid water content, precipitation type, and temperature as inputs. It also uses precipitation type and cloud type to determine whether the icing is continuous or intermittent. In the original LAPS framework, values of the LWC were determined from calculations based on the Smith-Feddes model (as described in Albers et al. 1996) applied to the entire three-dimensional database. For ARPS forecasts complete three-dimensional water mixing ratio profiles already exist and can be used to compute a LWC from the summation of the mixing ratios, q_c and q_r. The LWC (g m^{-3}) is determined by the following conversion:

$$LWC = \rho_d (q_c + q_r) \tag{3.23}$$

where ρ_d is the air density of dry air.

The precipitation type is determined using another adaptation of LAPS code based on the analysis of the wet bulb temperature (T_w). The five states of precipitation that can be assessed are rain, snow, ice pellets, freezing rain, or hail. The wet bulb temperature is calculated using an empirical integration based on algorithms for

computing selected meteorological quantities developed by Stipanuk (1973). For the entire domain, each column of model grid points is analyzed from the top down. Model derived water substance, radar reflectivity (Z; dBZ), is computed based on

$$Z = 17300(1000\rho q_r)^{1.75} + 38000(1000\rho(q_s + q_h))^{2.2} \qquad (3.24)$$

where ρ is the total air density, q_r is the rain water mixing ratio, q_s is the snow water mixing ratio, and q_h is the hail water mixing ratio (Kessler 1969). The column is then examined throughout the region(s) where reflectivity exists. Initially the precipitate is diagnosed as snow at the top of the reflectivity regions(s) if the T_w is less than 0°C; otherwise it is set to be rain. The precipitate is then tracked towards the surface and is continuously tested layer by layer within the reflectivity region to see how it changes. In a layer of T_w less than 0°C, frozen precipitation (snow or ice pellets) remains unchanged. Liquid precipitation freezes as ice pellets if the total area integrated vertically over pressure in the layer where $T_w < 0$ °C, is greater than 25,000 °C Pa. Otherwise, the liquid precipitation is classified as freezing rain. If the precipitate falls through a layer of T_w greater than 1.3°C, frozen precipitate and freezing rain are changed into rain. This threshold is set above 0 °C to take into account the time needed for the precipitate to melt. If precipitate falls through a layer of T_w between 0°C and 1.3°C, it remains unchanged. Finally, after the precipitate is tracked down to the surface, a second pass is performed. In areas with reflectivity greater than 50 dBZ, anywhere in the column, the precipitate is changed to hail. The algorithm results in a three-dimensional precipitation type gridded field. Recent ARPS modifications to output precipitation rates at the surface will be incorporated into updates to the algorithm to remove the need to rely on reflectivity based precipitation assessments.

Using the model temperature, LWC profiles, and precipitation type, the resulting icing severity forecast is derived from the thresholds in Table 3.4. A sample forecast product is shown in Figure 3.10. As before, the shaded areas indicate icing forecast with the text overlay showing the icing "frequency" forecast (continuous or intermittent).

Even though the cloud type assessment is only used to specify icing frequency conditions, which are not used for verification purposes, a brief description of how the cloud type assessment is made is appropriate. The algorithm uses the model forecast cloud water (q_c) and ice water (q_i) values as the cloud indicator when the sum exceeds a specified threshold, currently set at $q_c + q_i > 0.01$ g kg^{-1}. If clouds are present, evaluation of the temperature and vertical gradient of the equivalent potential temperature ($\partial\theta e/\partial z$; °K m^{-1}) characteristics results in the assignment of one of nine different cloud types. These include stratus (ST), stratocumulus (SC), cumulus (CU), cumulonimbus (CB), altostratus (AS), altocumulus (AC), cirrus (CI), cirrostratus (CS), and cirrocumulus (CC). The assignment of cloud types is according to Table 3.5.

Temperature, °C	$\partial\theta e/\partial z$ (°K m^{-1}) Threshold and Cloud Type Assignment			
T > -10 °C	$\partial\theta e/\partial z > 0.001$ **ST**	$-0.001 < \partial\theta e/\partial z < 0.001$ **SC**	$-0.005 < \partial\theta e/\partial z < -0.001$ **CU**	$\partial\theta e/\partial z < -0.005$ **CU** (if vertical extent > 5km **CB**)
-20 < T < -10 °C	$\partial\theta e/\partial z > 0$ **AS**	$\partial\theta e/\partial z < 0$ **AC**		
T < -20 °C	$\partial\theta e/\partial z > 0.0005$ **CS**	$-0.0005 < \partial\theta e/\partial z < 0.0005$ **CI**	$\partial\theta e/\partial z < -0.0005$ **CC**	

Table 3.5: Overview of LAPS Cloud Type Algorithm based on temperature (T; °C) Vertical Gradient of the Equivalent Potential Temperature ($\partial\theta e/\partial z$; °K m^{-1}) thresholds.

Icing Severity	Liquid Water Content (LWC), Temperature (°C), and precipitation type thresholds
Light	$0.01 < LWC \leq 0.1$ **and** $T \leq -10$
Moderate	$0.1 < LWC \leq 0.5$ **and** $T \leq -10$ or $0.01 < LWC \leq 0.5$ **and** $-10 < T \leq -5$ or $0.01 < LWC \leq 0.1$ **and** $-5 < T \leq 0$
Heavy	$LWC \geq 0.5$ **and** $T \leq 0$ or $0.1 < LWC \leq 0.5$ **and** $-5 < T \leq -0$ or Freezing Rain is assessed at grid point

(a)

Icing Frequency	Cloud and Precipitation Types
Continuous	Stratus, nimbostratus, altostratus, cirrostratus, or cirrus clouds, or freezing rain
Intermittent	Stratocumulus, cumulus, altocumulus, cirrocumulus, or cumulonimbus.

(b)

Table 3.4: Overview of LAPS (a) Icing Severity Index and (b) Icing Frequency Algorithm based on temperature (T; °C), Liquid Water Content (LWC; g m^{-3}), cloud type, and precipitation thresholds. (Albers; personal communication)

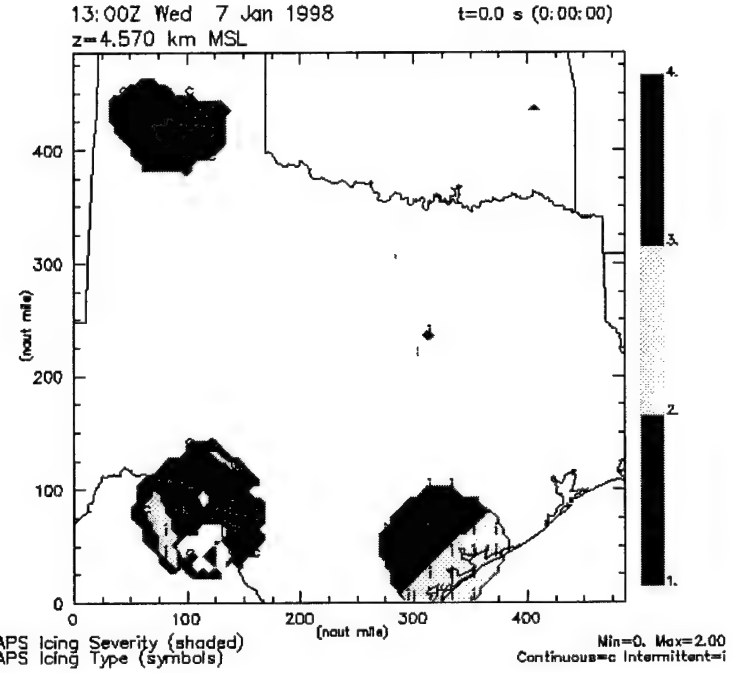


Figure 3.10: Sample LAPS Icing Algorithm forecast product. Existence of icing indicated by the color shading (1 – Light, 2 – Moderate, 3 – Heavy), frequency indicated by text (c – Continuous and i – Intermittent).

(5) *The STOVEPIPE Icing Algorithm*

The fifth threshold-based algorithm is the STOVEPIPE Algorithm (Bernstein 1996), also developed by NCAR/RAP. It derives its name from the way in which it assesses icing characteristics in a specified vertical column of a three-dimensional forecast or analysis database. It is an advancement of the previous NCAR/RAP algorithm in that it attempts to remove regions of icing meeting broad temperature and relative humidity criteria by introducing the occurrence of specific precipitation types and cloud cover amount thresholds.

The algorithm as established by Bernstein used actual surface observations of cloud cover and precipitation type as input into the routine. At each horizontal model or analysis grid point, all surface observations within a range of influence are examined to determine a worst case cloud cover amount (whether the sky is overcast or not) and precipitation type if present. If certain surface weather conditions exist, at a grid point, the algorithm then examines the vertical column for areas which meet specified temperature and relative humidity thresholds. For forecasts, it was proposed that the most current observed conditions available could be used as a "persistence" forecast. This, however, could lead to erroneous areas of icing because of the movement of weather features. For the forecast icing assessments needed by COMET-Tinker, forecast surface precipitation type and cloud cover amounts must be derived from separate algorithms as a substitute for the observational data.

The forecast element requirements of the COMET-Tinker project from the initial Tinker AFB inputs included the surface precipitation type. Two precipitation type algorithms are included in the real-time AIV production suite, one based on the current

Meso-ETA model (Baldwin and Contorno 1993) and one based on the LAPS code previously discussed. Preliminary evaluation of the resultant surface forecasts showed significant agreement between the algorithms in the precipitation type estimate and their corresponding spatial coverages. To maintain a consistent precipitation type field for both the Stovepipe and LAPS icing determinations, the LAPS precipitation type assessment was used as input into the Stovepipe algorithm. For the cloud cover amount a simple dew point depression examination, illustrated in Table 3.6 was used to specify the cloud cover amount as viewed from the surface. This assumed no overlap in cloud regions. Therefore the cloud cover amount at the vertical level that exhibited the greatest coverage was assigned as the "surface" cloud cover amount used in the algorithm.

Dew point Depression °C	Surface Cloud Cover Category
< 2	Overcast (OVC)
2 to 3	Scattered to Broken (SCT- BKN)
3 to 4	Scattered (SCT)
4 to 5	Scattered to Few (SCT-FEW)
> 5	Clear (CLR)

Table 3.6: Cloud cover amount assessments used in the STOVEPIPE Icing Algorithm based on dew point depression (DD; °C) thresholds.

Through the evaluation of over 3500 icing PIREPs, Bernstien (1996) established an empirical database identifying the related surface weather and vertical temperature and relative humidity structure where icing conditions were observed. Of the full set of PIREPs only those associated with moderate or greater icing intensity, called "WORST PIREPs", were used to isolate those conditions which have the greatest probability of being associated with icing of significant interest to the flying community. This restriction reduced the number of icing reports to roughly 10% of the original amount (Bernstein 1996).

The observed surface cloud cover amounts and precipitation types within a range of influence of 160 km (250 km when west of 100°W longitude) of reported icing areas were determined. The precipitation types were broken down into the following categories: drizzle (L), rain (R), snow (S), ice pellets/sleet (IP), freezing drizzle (ZL), and freezing rain (ZR). The cloud cover categories used were obscured (XOB), overcast (OVC), broken (BKN), scattered (SCT), and clear (CLR) sky. In cases with multiple observed conditions, either precipitation and/or cloud cover, within the specified range of influence, the icing reports were associated with each condition experienced separately. Applying all observed surface conditions in the range of influence has the drawback that the occurrence of snow and rain is much more prevalent than the freezing precipitation types and we would expect them to correspond to a larger number of reported icing areas. To adjust the relationships derived for the differences in the spatial coverage of the various observed categories, an icing "Threat" given by

$$\text{Threat} = \frac{\text{number of "WORST" icing PIREPs}}{\text{Area extent of precipitation or cloud cover category}} \tag{3.25}$$

which measuring the average number of observed icing reports per unit area (km^2) of each precipitation type or cloud cover amount category.

Nearly 67 % of the "WORST" PIREPs were associated with some form of surface precipitation. While only a third of those were associated with freezing precipitation types (ZR and ZL), they produced the greatest threat score, illustrated in Figure 3.11. This was mainly due to the extensive spatial coverage of rain and snow cover versus the much smaller regions with freezing precipitation in the observed database. Thus an

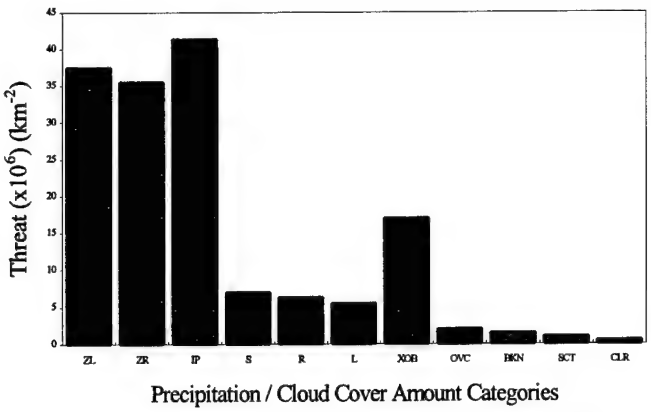


Figure 3.11: Threat (# of WORST PIREPs per unit area; x 10^6) (km^{-2}) calculated for each precipitation type and cloud cover amount category. Category abbreviations follow standard contractions (e.g. ZL is freezing rain, XOB is obscured sky cover, etc) (after Bernstein 1996)

association of icing reports with freezing precipitation can be extracted, which agrees with the intuitive expectation of the existence of some amount of supercooled liquid water (SLW) in the column above where freezing precipitation is being observed. For these SLW icing cases, an evaluation of the temperature and relative humidity profiles around the flight level where the icing was reported showed that most of the ZR, ZL, L, and IP events occurred when $-12\ °C \leq T < 0\ °C$ and $RH > 75\%$ (Bernstein, personal communication). The first part of the Stovepipe algorithm, shown in overview in Table 3.7, is designed to look for these indications of supercooled liquid water in the forecast domain.

Precipitation Type or Cloud Cover Amount	Freezing drizzle, freezing rain, or ice pellets at surface	Any precipitation, obscured, or overcast cloud cover at surface
Temperature, °C	$-12 \leq T \leq 0$ °C	$-15 \leq T \leq 0$ °C
Relative Humidity, %	≥ 75 %	$\geq 70\%$
Icing Forecast	Supercooled large drops (SLD)	General Icing

Table 3.7: STOVEPIPE Icing Algorithm based on temperature (T; °C), relative humidity (%), precipitation type, and cloud cover amount thresholds.

The remaining 33% of the WORST PIREPs were not associated with any form of observed precipitation. The majority of those were in regions of either BKN or OVC sky conditions. As the amount of cloud cover decreased, a corresponding decrease in the number of observed icing reports was evident. For these cases, as well as the S and R precipitation cases, a broader set of thresholds, $-15\ °C \leq T < 0\ °C$ and $RH > 70\%$ was used. Thus the second portion of the Stovepipe icing algorithm is geared to catch the non-freezing precipitating and general icing occurrences relationships which were found to exist .

The STOVEPIPE algorithm first performs the check (labeled SLD in Table 3.7) for the freezing precipitation or ice pellets (sleet) at the surface based on the two-dimensional surface precipitation type forecast from the applicable (LAPS) algorithm. If present, the SLD temperature and humidity thresholds are evaluated at each grid point in the vertical above that surface point. Since the SLD check is designed to catch supercooled liquid water, if a $T > 0\ °C$ region, beginning more than 30m above the ground, is detected above a sub-freezing grid point, all grid points above the level of the warm layer are not examined for the SLD conditions. This is due to the assumption that the freezing precipitation forecast was caused by the melting of frozen precipitates falling through the warm layer. In other words, no significant amount of "liquid water" is assumed to exist above the warm layer which could lead to icing conditions. For all grid points not meeting the SLD conditions, the general icing check is then performed.

Since the COMET-Tinker version of the algorithm uses a forecast derived precipitation type it is considered a "Forecast–Based" version. The main drawback to the "Forecast–Based" nature of the algorithm is that it could introduce areas of potential

icing that would be eliminated from a purely "observational" assessment. This is because all grid points have some type of "likely precipitation" determined whether or not the model actually forecast the precipitation to occur. Thus all horizontal grid points could be checked by the algorithm. The concept of removing non-precipitating area through the use of forecast radar reflectivity or other inherent model variable was used. During WOP97 the code was set up to use the presence of forecast radar reflectivity at the lowest level (k=2) to remove areas with no "forecast" precipitation. As with the Stovepipe algorithm, recent additions of rain rate calculations to the ARPS model code may lead to easier determinations of whether there is actual precipitation being forecast. Figure 3.12 shows a sample product derived from the Stovepipe algorithm. Areas with a YES forecast are shaded with the appropriate formation mechanism denoted by symbols.

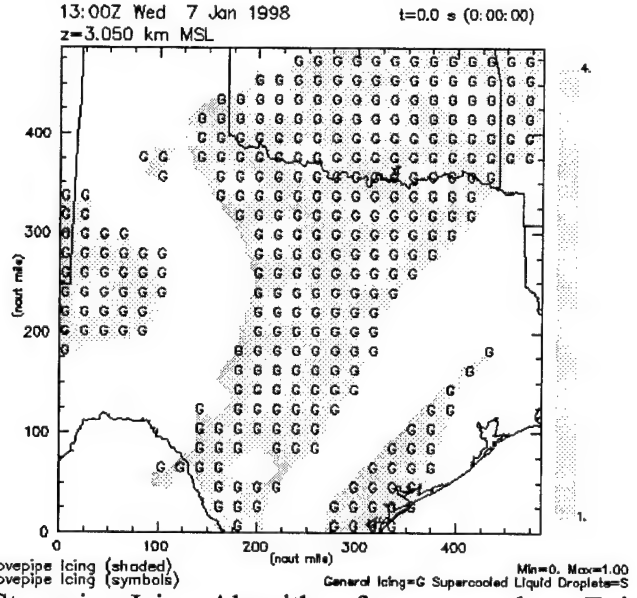


Figure 3.12: Sample Stovepipe Icing Algorithm forecast product. Existence of icing indicated by the color shading, type indicated by text (G – General Icing, S – Supercooled Liquid Droplets).

(b) *Cloud Microphysics Parameterization Based Icing Algorithms*

The threshold-based algorithms previously discussed primarily use empirical formulations to attempt to diagnose regions of "cold clouds". Originally designed for upper-air sounding evaluations completed by forecasting personnel, they have been adapted to be used on numerical model output. Interest in higher resolution mesoscale forecasting models such as the Pennsylvania State University/NCAR Mesoscale Model 5 (MM5) and ARPS has led to the need to handle model processes on an increasingly smaller scale. The ability of these models to computationally assess microphysical processes in clouds through detailed parameterization schemes has enabled model output to include direct evaluation of liquid water content, hydrometeors, small scale vertical velocities, and other quantities which may yield improved forecasts of hazardous weather conditions to the aviation community. The second type of icing algorithm that has been included in this work is a result of direct computations using microphysical parameterization schemes within mesoscale forecast models.

(1) *Tremblay Supercooled Liquid Water Icing Algorithm*

The first algorithm to assess the presence of supercooled liquid clouds that is based on microphysical parameterization schemes is from studies by Tremblay et al. (1995; 1996) of the Cloud Physics Research Division of the Atmospheric Environment Service, Canada. It works by associating the generation of supercooled liquid water to the potential for in-flight icing conditions to exist.

As part of the Canadian Atlantic Storms Program (CASP), aircraft flights were made through a multitude of different weather-producing systems. Two aircraft flights (5 and 29 February 1992) were made through systems in which a significant amount of

supercooled liquid water was experienced. These two days were isolated for comparison with numerical simulations of the storms using the Canadian operational regional finite element (RFE) model, which included parameterizations of cloud microphysics. The RFE version evaluated used parameterizations developed by Sundqvist et al. (1989) (as described in Tremblay et al. 1996); however ice microphysics were not included.

As a result of the simulations, the most significant microphysical processes within clouds composed of supercooled liquid water were identified. These processes primarily involved the generation of supersaturated water vapor by wet adiabatic cooling, wG, where w is the vertical velocity and G is the vertical gradient of the saturation mixing ratio ($\partial w_s/\partial z$). Of the thirty-nine (39) processes present in the parameterizations, only the seven (7) shown in Figure 3.13, normalized against wG, exhibited any significance (with a value greater than one on the normalized scale). The acronyms for the processes are described in Table 3.8.

From Figure 3.13. the following simplified model for microphysics inside steady-state supercooled clouds was proposed:

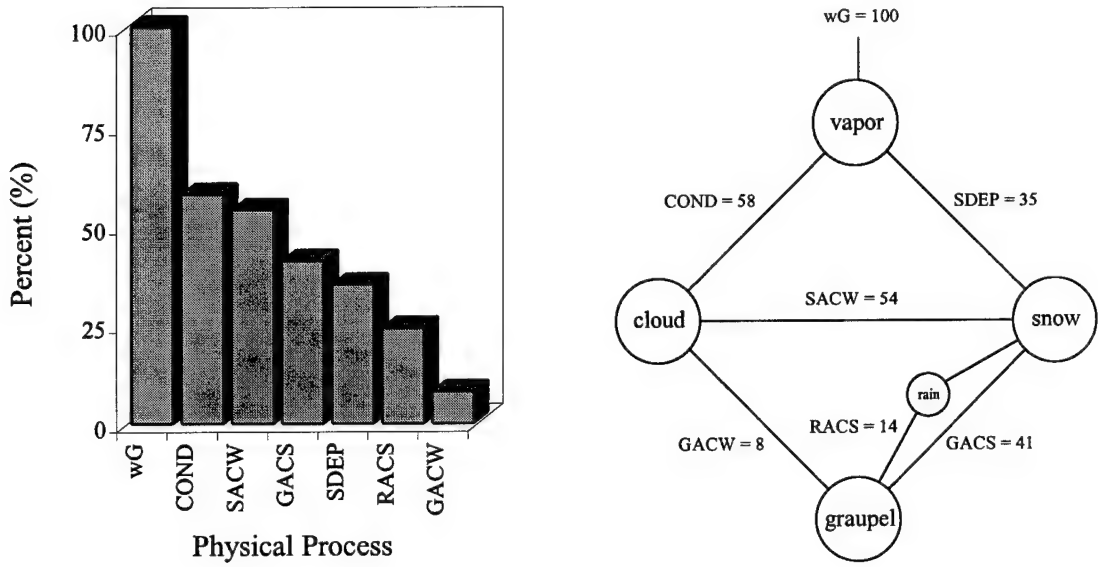
$$\text{COND} \approx (\text{wG} - \text{SDEP}) \approx \text{SACW} \tag{3.26}$$

Through empirical relationships, SACW is related to the amount of cloud water (q_c) by

$$SACW = \left[\frac{\pi}{4}\rho \mathrm{E}_{sc} a M_s^{2+b}\left(\frac{\rho_o}{\rho}\right)^{1/2}\right] q_c \equiv [\Phi] q_c \tag{3.27}$$

where ρ is the air density, ρ_o is a reference air density, E_{sc} is the collection efficiency of cloud water by snow , M_s^J is the J^{th} moment of the snow distribution, and a and b are an empirically based constant, we can combine 3.26 and 3.27 to get

$$q_c = \frac{wG - SDEP}{\Phi} \tag{3.28}$$



(a) (b)

Figure 3.13: Interactions of significant microphysical processes within supercooled clouds. (a) Relative rate of the production of meteors, normalized by wG (b) simplified model interactions. Number refers to (a) (Abbreviations listed in Table 3.8) (after Tremblay et al. 1996)

Notation	***Description***
wG	generation of supersaturated water vapor
COND	condensation of water vapor
SACW	collection of cloud water by snow
GACS	collection of snow by graupel
SDEP	deposition on snow
RACS	collection of snow by rain to produce graupel
GACW	collection of cloud water by graupel

Table 3.8: Nomenclature for predominant microphysical mechanisms that generate supercooled liquid water.

Thus there will be a net generation of q_c if the difference (wG – SDEP), the excess amount of water vapor being produced that cannot be converted into snow, is larger than zero. Vertical velocity, w, is a model variable and can used directly. The vertical gradient of the saturation mixing ratio, G, can be expressed using the Clausius-Clapeyron equation as

$$G = \rho r_s \left[\frac{L_v}{R_v T^2} \Gamma_w(T,p) - \frac{g}{R_d T} \right] \tag{3.29}$$

where r_s is the saturation mixing ratio, L_v is the latent heat of vaporization, R_v and R_d are the respective gas constants for water vapor and dry air, T is the air temperature, p is the pressure, g is the acceleration due to gravity, and Γ_w is the moist adiabatic lapse rate. Finally, if it is assumed that supercooled clouds are saturated with respect to water, the deposition on snow, SDEP, can be determined from integrating the crystal growth rate over the entire snow size distribution to yield

$$SDEP = \left(\frac{2\pi [S_i - 1]}{\frac{L_s^2}{K R_v T^2} + \frac{R_v T}{e_{si} D}} \right) f \times Ms^{(1)} \tag{3.30}$$

where $S_i \equiv e_s / e_{si}$ the ratio of the saturation vapor pressures over water and ice, L_s is the latent heat of saturation, K is the coefficient of thermal conductivity of air, D is the coefficient of diffusion of water vapor in air, f is a dimensionless adjustment factor, $Ms^{(1)}$ is the first moment of the snow particle size distribution and is empirically related to the snow distribution, q_s, by equation 3.30.

$$Ms^{(1)} = 55.6 \left(\frac{6 q_s}{\pi \rho_s} \right)^{0.58} \tag{3.31}$$

The value of f needs to be empirically determined to account for the inadequacies in the snow size distribution parameterizations and the nature of the vertical velocity field in the model. These features are highly dependent on the resolution of the model. For the simulations conducted by Tremblay et al. (1996), a value of f = 0.1 was determined. However, the implementation of the code into the COMET-Tinker display software f was reset to 1.0. Additional testing with the algorithm will be required to determine the best choice of f for various ARPS model resolutions, especially if there is a noticeable underforecasting of icing by the Tremblay algorithm. Figure 3.14 shows a sample display of the Tremblay icing algorithm product available to the forecasters. The plotted

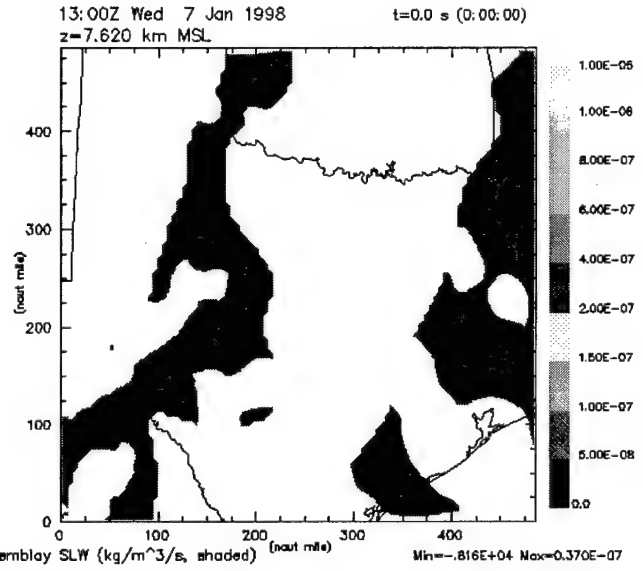


Figure 3.14: Sample Tremblay SLW Icing Algorithm forecast product. Existence of icing indicated by any color shading (wG – SDEP > 0).

variable, (wG – SDEP), is adjusted to units of kg m^{-3} s^{-1} to show the SLW generation rate if the air temperature is between 0 and –40 °C. It is shaded for various values that are greater than zero to indicate icing potential. While no specific relationship between icing intensity and the amount of SLW production can be inferred from the plot, areas

with greater values can be reasoned to have a higher likelihood of icing than regions with production at or near zero.

(2) *ARPS Supercooled Liquid Water Icing Algorithm*

The second icing algorithm of this type utilizes the parameterized microphysics that are contained in the ARPS forecast model (Xue et al. 1995). It is the only one in which we see the direct application of meso-scale numerical liquid water content calculations to the evaluation of icing.

The general conservation equation for the mixing ratios of water vapor (q_v), cloud water (q_c), rain water (q_r), cloud ice (q_i), snow (q_s), and hail/graupel (q_h) can be expressed for a summed mixing ratio variable q_ψ as

$$\frac{\partial}{\partial t}\left(\rho^* q_\psi\right) = -\underbrace{\left[\rho^* u \frac{\partial(q_\psi)}{\partial \bar{\xi}} + \rho^* u \frac{\partial(q_\psi)}{\partial \bar{\eta}} + \rho^* W^c \frac{\partial(q_\psi)}{\partial \bar{\zeta}}\right]}_{①} + \underbrace{\frac{\partial\left(\rho^* V_{q_\psi} q_\psi\right)}{\partial \bar{\zeta}}}_{②} + \underbrace{\sqrt{G} D q_\psi}_{③} + \underbrace{\sqrt{G} S q_\psi}_{④} \qquad (3.32)$$

with ① being the advection term, ② the sedimentation (i.e. falling out of rain, snow and, hail at their terminal velocity), ③ mixing, and ④ the source/sink term. In 3.32, ρ is the air density; ξ, η, and ζ are unit vectors of the curvilinear coordinate system used by ARPS, u is the Cartesian x-direction velocity component, W^c is the covariant vertical velocity component, Vq_ψ is the terminal velocity of the respective water states, $\sqrt{G}$ is the determinant of the Jacobian matrix transformation from the curvilinear (ξ, η, ζ) coordinate system to the Cartesian (x, y, z) given by

$$\sqrt{G} \equiv \frac{\partial(x,y,z)}{\partial(\xi,\eta,\zeta)} = \begin{vmatrix} x_\xi & x_\eta & x_\zeta \\ y_\xi & y_\eta & y_\zeta \\ z_\xi & z_\eta & z_\zeta \end{vmatrix} \qquad (3.33)$$

Finally, Dq_ψ and Sq_ψ are the summed diffusion and source/sink terms, respectively, which include all parameterized microphysical processes. The diffusion is accomplished by either a second and/or fourth order computational mixing scheme. The source/sink is usually handled either by a microphysics package modeled after a scheme developed by Kessler (as described in Xue et al. 1995) which only includes "warm rain" microphysics parameterizations, or one after Lin-Tao (as described in Xue et al. 1995).

The Lin-Tao scheme is the most complete microphysics scheme within the ARPS code and was used exclusively during the winter operational period. It includes the Kessler microphysics and a three-category ice phase parameterization scheme initially developed by Lin (1983) (as described in Xue et al. 1995). The three ice phases include cloud ice (q_c), snow (q_s), and hail/graupel (q_h). The overall interactions between the various components are illustrated in Figure 3.15., with symbol descriptions in Table 3.9.

The Kessler warm-rain parameterization scheme in ARPS has the ability to modify three separate water states within the model physics: water vapor, cloud water, and rain water. The following discussion treats the Kessler scheme separately, as if it were the only microphysics scheme being treated by the model. Adjustment to the three categories occurs under the following conditions: (a) Autoconversion of cloud water to rain water approximated by

$$A_r = C_{ar}\left(q_c - q_{ccrit}\right) \tag{3.34}$$

where A_r is the autoconversion rate (kg kg^{-1} s^{-1}), q_c is the cloud water mixing ratio (kg kg^{-1}), and $q_{c\ crit}$ is the cloud water mixing ratio threshold (kg kg^{-1}). (C_{ar} is a constant determining the rate of conversion, s^{-1}); (b) Accretion (collection) of cloud water by

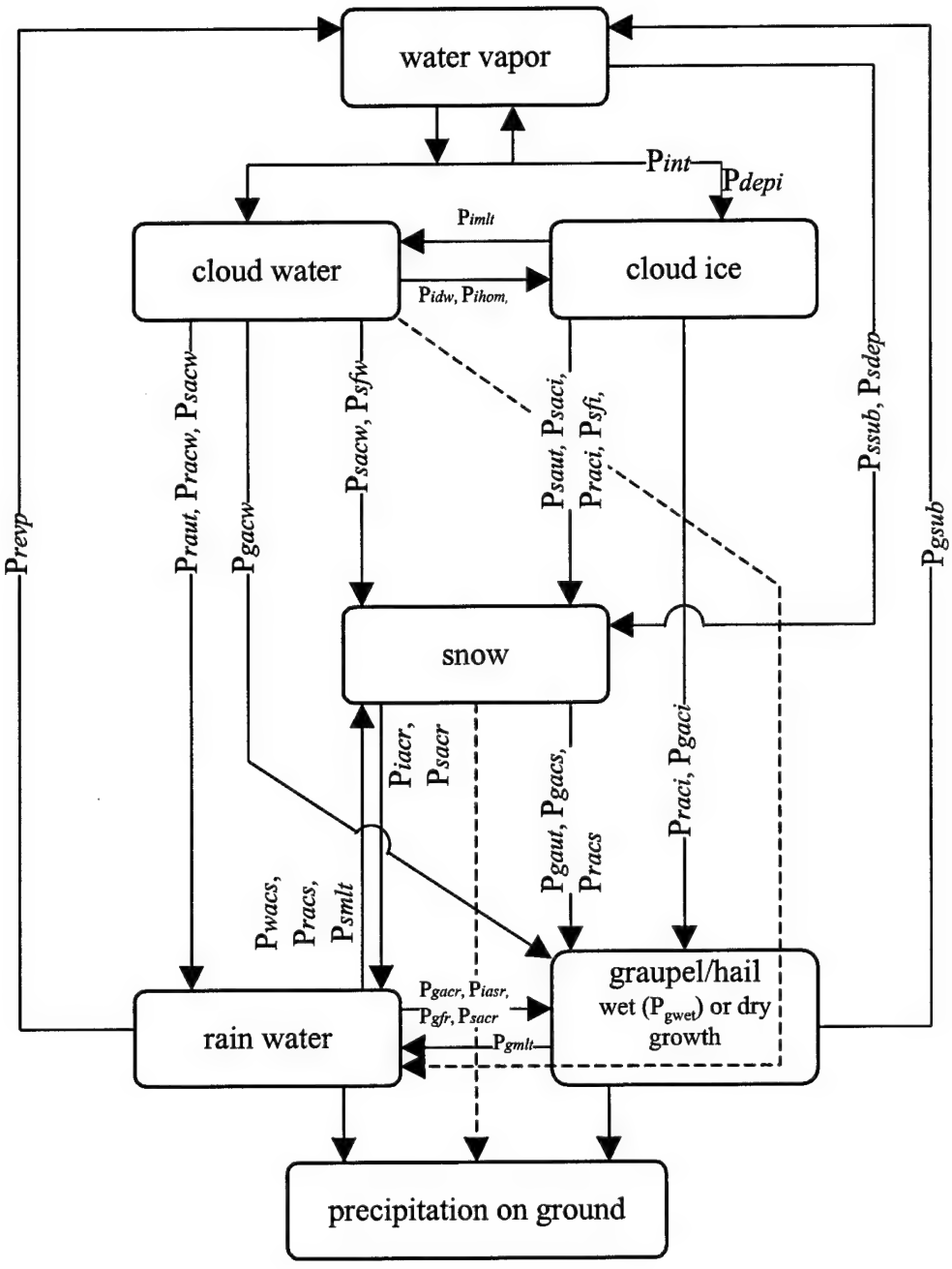


Figure 3.15: Cloud microphysical processes considered in the ice microphysics parameterization scheme (after Xue et al.1995; Lin et al. 1983)

Symbol	Definition
P*depi*	Depositional growth of cloud ice
P*int*	initiation of cloud ice
P*imlt*	melting of cloud ice to cloud water
P*idw*	depositional growth of cloud ice at the expense of cloud water
P*ihom*	homogeneous freezing of cloud water to cloud ice
P*iacr*	accretion of rain by cloud ice, producing snow or graupel, depends on amount of ice
P*raci*	accretion of rain by cloud rain, producing snow or graupel, depends on amount of rain
P*raut*	autoconversion of cloud water to rain
P*racw*	accretion of cloud water by rain
P*revp*(e_r)	evaporation of rain
P*racs*	accretion of snow by rain, produces graupel or snow depending on temperature
P(Q)*sacw*	accretion of cloud water by snow, produces graupel or snow depending on temperature
P*sacr*	accretion of rain by snow, produces graupel or snow depending on amount or rain/snow
P*saci*	accretion of cloud ice by snow
P*saut*	autoconversion (aggregation) of cloud ice to snow
P*sfw*	Bergeron process (deposition and riming) – transfer of cloud water to snow
P*sfi*	Bergeron process embryos (cloud ice) – used to calculate transfer of cloud water to snow
P*sdep*(d_s)	deposition growth of snow
P*ssub*(S_s)	sublimation of snow
P*smlt*(m_s)	melting of snow to rain, T > 273.16 °K
P*wacs*	accretion of snow by cloud water to form rain, T > 273.16 °K
P*gaut*	autoconversion (aggregation) of snow to graupel
P*gfr*(f_g)	probabilistic freezing of rain to graupel
D(Q)*gacw*	accretion of cloud water by graupel
D(W)*gaci*	accretion of cloud ice by graupel
D(W)*gacr*	accretion of rain by graupel
P*gsub*(S_g)	sublimation of graupel
P*gmlt*(m_g)	melting of graupel to form rain, T > 273.16 °K
P*gwet*	wet growth of graupel

Table 3.9: Definitions of symbols in the ice microphysics parameterization scheme (after Xue et al. 1995)

$$C_r = C_{cr} q_c q_r^{0.875} \tag{3.35}$$

where C_r is the accretion rate (kg kg^{-1} s^{-1}) and q_r is the rain water mixing ratio (kg kg^{-1}). (C_{cr} is a constant determining the rate of conversion, s^{-1}); (c) evaporation of rainwater and cloud water, only when the air is unsaturated defined by

$$E_r = \frac{1}{=}{\rho}\left(\frac{C\left(1 - q_v / q_{vs}\right)\left[\bar{\rho} q_r\right]^{0.525}}{2.03\text{x}10^4 + 9.58\text{x}10^6 / \left[\bar{p} q_{vs}\right]} \right) \tag{3.36}$$

where E_r is the evaporation rate (kg kg^{-1} s^{-1}), q_v is the water vapor mixing ratio in kg kg^{-1}, p is the pressure in Pa, q_{vs} is the water vapor saturation mixing ratio in kg kg^{-1}, and all over-barred quantities are functions of height only. C is a ventilation coefficient given by

$$C = 1.6 + 30.3922\left(\bar{\rho} q_r\right)^{0.2046} \tag{3.37}$$

and lastly (d) conversion of water vapor to cloud water when the air is super-saturated ($q_v > q_{vs}$) or from cloud water to water vapor if unsaturated ($q_v < q_{vs}$). The change between q_v is given by δq_{vs}, the amount of cloud mixing ratio that is evaporated or condensed, which is derived from

$$\delta q_{vs} = \frac{-\left(\overset{*}{q}_{v} - \overset{*}{q}_{vs}\right)}{1 + \dfrac{a_w (273.15 - b_w) \overset{*}{q}_{vs} \left(L_v / C_p\right)}{\left[\overset{*}{T} - b_w\right]^2}} \tag{3.38}$$

with the asterisked variables already being updated for advection, diffusion, filtering, and other forcing processes. The final determinations of q_r and q_c, are made by doing a time differencing step, with Δt the integration time step, shown by

$$q_r^{n+1} = q_r^{*n+1} + 2\Delta t\left(A_r + C_r - E_r\right) \tag{3.38}$$

$$q_c^{n+1} = q_c^{*n+1} - \delta q_{vs} - 2\Delta t\left(A_r + C_r\right) \tag{3.39}$$

Other variables such as potential temperature, and water vapor are adjusted in a similar manner but are not utilized in the algorithm to assess the icing forecast.

To evaluate the icing forecast based on the ARPS supercooled liquid water, the amount of cloud water (q_c) and rainwater (q_r) at each grid point is summed. This gives us a measure of the amount of "liquid" water at the given location. If the air temperature is below freezing (T < 0° C), and there is *liquid* water, as represented by the summed value, then "supercooled liquid water" (SLW) is assumed and icing is forecast at that grid point. Figure 3.16 shows a sample product with an ARPS SLW icing assessment.

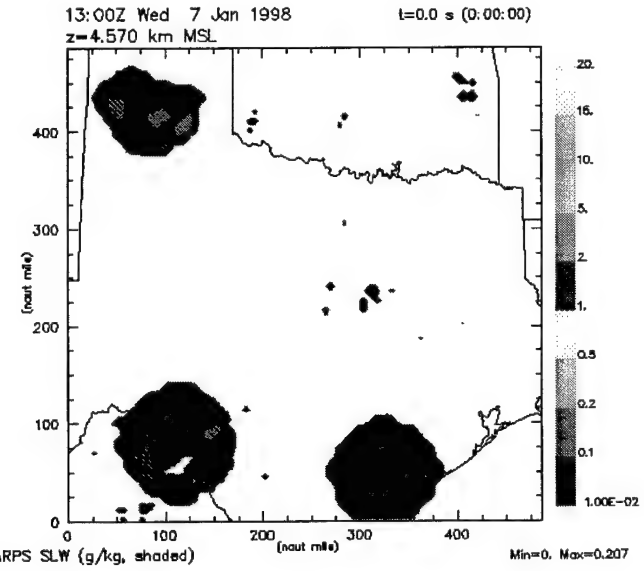


Figure 3.16: Sample ARPS SLW Icing Algorithm forecast product. Existence of icing indicated by color shading ($q_c + q_r$ mixing ratio > 0.01 g kg^{-1} with temperature < 0 ° C).

The plot shows the actual summed cloud and rainwater amount, $q_c + q_r$ in g kg^{-1}, for regions with temperatures below freezing. As with the Tremblay icing forecast,

empirically determined thresholds of the summed mixing ratio need to be determined to allow the diagnosis of an intensity or some relative form of "likelihood" of icing occurring.

Section 3.4.2 Icing Verification Processing

The process to accomplish the icing verification is a straight-forward computation similar to the surface and profiler data extraction routines. Using the above icing algorithms to produce a "Yes/No" gridded forecast, and the decoded PIREPs from the archive database with their "Yes/No" observations, the following procedures were performed.

The first task is to decide which PIREPs are applicable to each forecast output time. In other studies where forecast output was available relatively infrequently, on the order of every three to twelve hours, a large window of PIREPs valid over a particular period was desired to increase the sample size of the verification. This is considered reasonable due to the generally slow processes by which the upper atmospheric characteristics change. In many instances some PIREPs had to be disregarded because they were more than two to three hours from any model output time. However, the ARPS output is available at hourly intervals. For this reason, all PIREPs within thirty minutes (± 30 minutes) of the top of each hour were grouped together and compared with the appropriate icing forecasts for that hour.

The second step is to go inside the model to pull out the applicable forecast conditions for each observed report. The extraction of the "Yes/No" icing forecasts has two spatial procedures to be considered. The first must be to determine the model grid

point, or set of grid points, in the horizontal that most closely correspond to the observed location. The decoded location (latitude/longitude) of the report is used to determine the general location of the PIREP within the horizontal domain. The specific point(s) to use in the verification can then be evaluated based on one of two basic methods. In the first, the four grid points surrounding the report horizontally could all be incorporated the forecast icing assessment associated with the report. This method is called the "four point method". Alternately the "one point method" using just the nearest grid point could be used. Both can lead to incorrect assessments at the edge of regions in the model where there is icing forecast. However, in studies by Carriere et al. (1997) and Brown et al. (1997), the differences in verification results between the two methods were considered minimal. While direct interpolation to the actual latitude/longitude of the report can be used, the ability to adequately interpolate a "categorical forecast" makes this undesirable. For the verification done in this work the "one point" method was employed for computational simplicity.

Once the horizontal location within the model grid is identified, the second consideration is to find the vertical grid location(s) which match the altitude(s) of the PIREP. Again, there are two standard approaches to doing this. In the first case, the grid point nearest the actual flight level is used in the verification. In this manner, a simple one forecast to one observation comparison could be made. However, for most icing reports a range where the event was observed is indicated. To handle this, the second approach uses all grid points between those two altitudes to come up the icing forecast. A combination of these two methods is currently used. It is felt that the inclusion of a specific icing region in the report should be fully incorporated into the verification

procedure. For this reason, all vertical grid points in the indicated icing region are examined. It also includes a two thousand foot window above and below the reported region, to ensure the likelihood of more than one grid point in the region for comparison. As a backup, a comparison at the grid point nearest the reported flight level is used as a means to break ties. This occurs when the icing region to be examined includes an even number of grid points that indicate equal numbers of correct and incorrect forecast points. For example if 8 grid points are examined, with 4 indicating a "Yes" icing forecast while the other 4 have a "No" forecast, the grid point closest the flight level will be used as the icing assessment for that comparison. While the flight level doesn't always necessarily match the indicated icing region, for the majority of PIREPs it is relatively close and therefore may have a better indication, than the other gird points examined, of the quality of the forecast associated with that report. The inclusion of multiple grid points, both horizontally (i.e. four point method) and vertically (i.e. entire range of icing reported), in the comparison with each PIREP is designed to provide a "smoothing" mechanism on the icing regions forecast. To restrict the verification to a single model grid point would severely punish the model. The presence of even the slightest phase error would make the forecast appear to have no ability in predicting icing.

The resulting forecast/observation pairs are then compiled for each individual forecast hour, for each daily forecast run, and the entire operational period as a whole. The last step in the PIREP verification is to sum the area (volume) associated with each icing algorithm forecast to determine the total impacted area (volume) characteristics of the icing regions. A separate calculation is done for each individual algorithm that is used in the forecast suite.

Chapter 4: Discussion of Results

Forecasts of icing aviation impact variables are among the core operational requirements for Air Force weather forecasting. This type of forecast product has uses within the generation of the base terminal forecast, development of warning and resource protection services, and flight weather briefing support provided to aircraft pilots. The ability of a numerical prediction system, of any size and resolution, to accurately assess the locations and intensities of in-flight icing is critical to the quality and usefulness of model-based forecast products provided to a BWS. The previous discussion of real-time forecasting operations conducted by COMET-Tinker, the development and implementation of several icing algorithms from the forecast output, verification techniques and procedures, and the observational data collection that was performed, give us a framework to assess the relative usefulness of the icing forecasts produced during WOP97.

The discussion of verification results will be divided into two parts. In the first part, the relationship between the observed meteorological variables at the surface and upper-air, and the model forecasts of these variables will be used to provide an overall quality assessment of the forecasts. This will indicate whether the comparison of icing forecasts based on algorithms using these forecast variables is worth the effort. The desire is to ensure that the model variables and parameters required by the algorithms making the icing forecasts do not have gross systematic errors. A complete understanding of the nature of these errors will not be discussed and would be handled better by case study analysis and more rigorous evaluation of individual forecasts. If

forecasts of the primary variables are considered reasonable, the second step is to examine the resulting ability of each algorithm to forecast icing conditions as compared to the pilot reports transmitted during the forecast period. It is hoped that a relative effectiveness ranking of each algorithm can be determined. The methodology and physical rationale behind each algorithm can then be used to attempt to explain its performance.

Section 4.1 Verification of Surface Forecasts

It has been demonstrated that surface flux and boundary layer issues, treatment of initial soil conditions, surface pressure forecasts and a related wind bias, and dependency on physical parameterizations which are still undergoing improvements can all lead to errors in the ARPS forecast model (Mewes 1997). In verification of the surface forecast conditions by ARPS, forecast errors were noted, such as a warm temperature biases, initialization errors, reduced wind speed indications, and significant high moisture (dew point) predictions. These agree with forecast errors identified in previous verification results (Mewes 1997). The average bias and RMSE measures of all surface station forecasts averaged over all forecast times is shown in Figure 4.1. Temperature (°C), dew point (°C), wind speed (ms^{-1}) and direction (°), station pressure (mb), and equivalent potential temperature (K) are plotted in a surface meteogram format. The convention used for *wind direction* errors is such that a positive error indicates a forecast wind direction turning clockwise from the observed (backing) and the opposite (veering) for a negative error in direction.

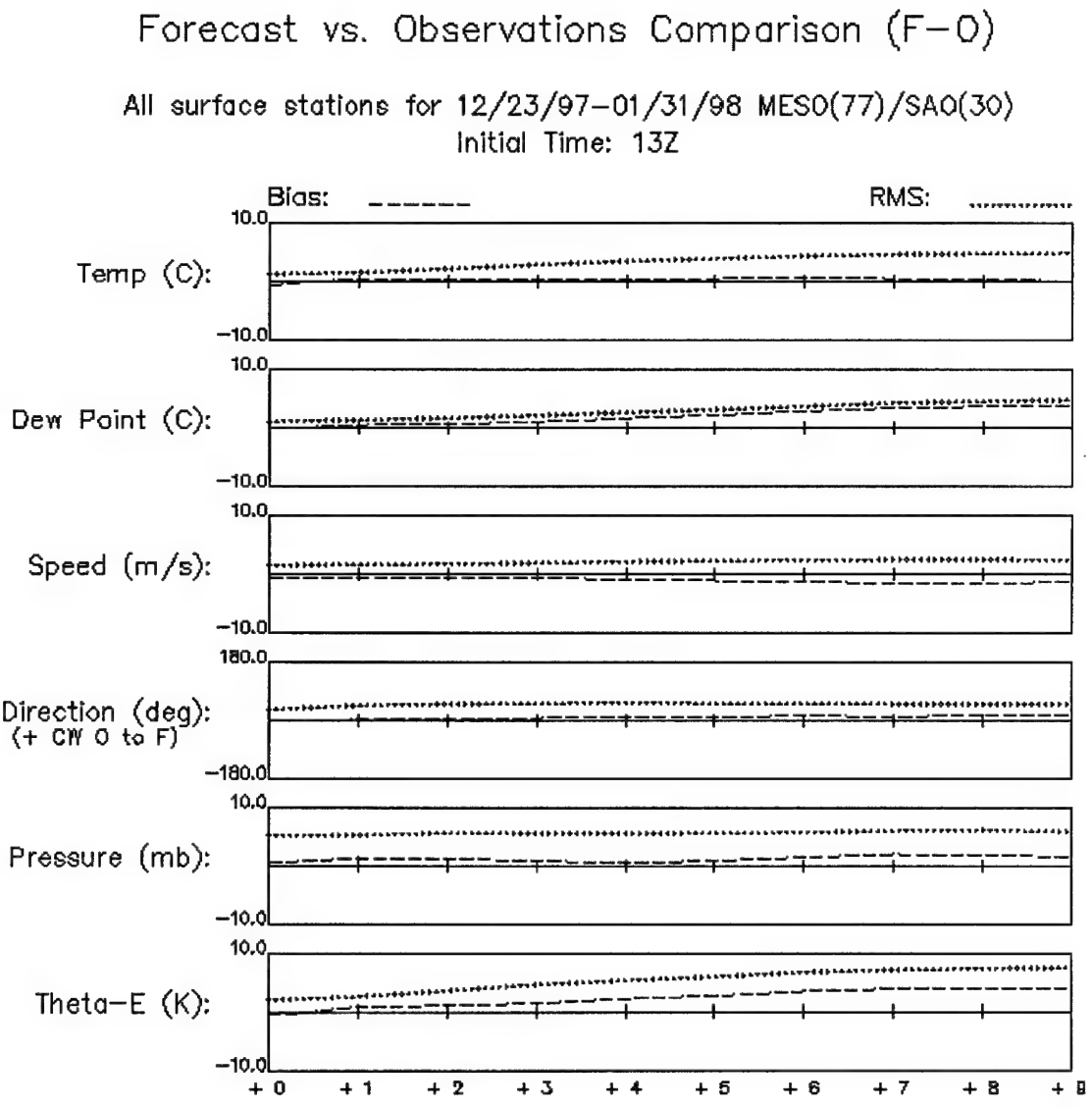


Figure 4.1: Average bias and RMSE values, for each valid hour, over all surface stations and all daily WOP97 forecasts combined.

It is desired that the bias from a long series of daily forecasts should be near zero. This is, however, not true with the moisture forecasts at the surface. The average biases show a consistent "moist" bias in the ARPS model at the surface. All parameters (except dew point) have RMSE which are significantly larger then the corresponding biases. Therefore, even though some appear accurate from a raw bias standpoint, each individual forecast may still have significant errors of both positive and negative direction. This is illustrated in Figures 4.2 and 4.3. The forecast results averaged over 24 December 1997 alone show agreement with expected ARPS forecast errors. A warm bias, increasing with time, is present as is a positive moisture bias. The opposite is true for the daily average over the 27 January 1998 forecast which shows an increasing negative temperature bias (i.e. "cold" bias) in the model forecast. The similarity of the bias with respect to the RMSE values shows that the forecast is consistently "cold" throughout the entire forecast domain.

A complete summary of the daily temperature bias over all surface stations is presented in Figure 4.4. To make the chart easier to read, only forecasts valid at three hour intervals (00, 03, 06 and 09 hours; 13, 16, 19, and 22 UTC respectively) are plotted. Here we see the marked variability in the overall sign of the forecast error showing both days with a warm bias and days with a cold bias. In all instances we see a consistent increase in the overall magnitude in the error as the forecast proceeds and in most cases begins to increase relatively early in the forecast period. All days have an increasing RMSE error with time shown in Figure 4.5. Again, we see that for the "warm" bias days the RMSE increases steadily with time, while the increase on "cold" bias days is mostly early in the forecast and levels off with time. One consistent feature

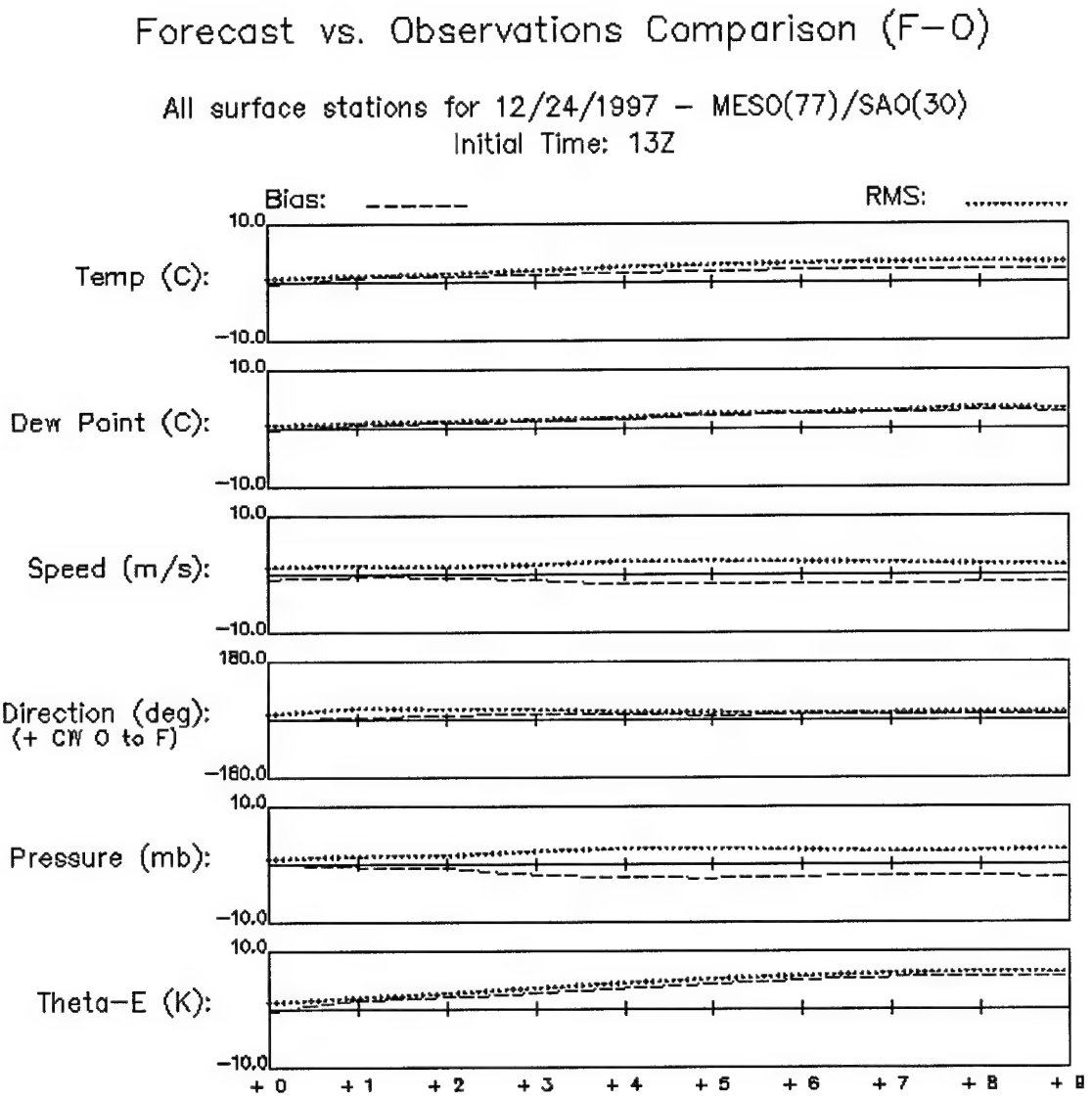


Figure 4.2: Same as Figure 4.1 except only averaged over surface stations for the forecast made on 24 December 1997.

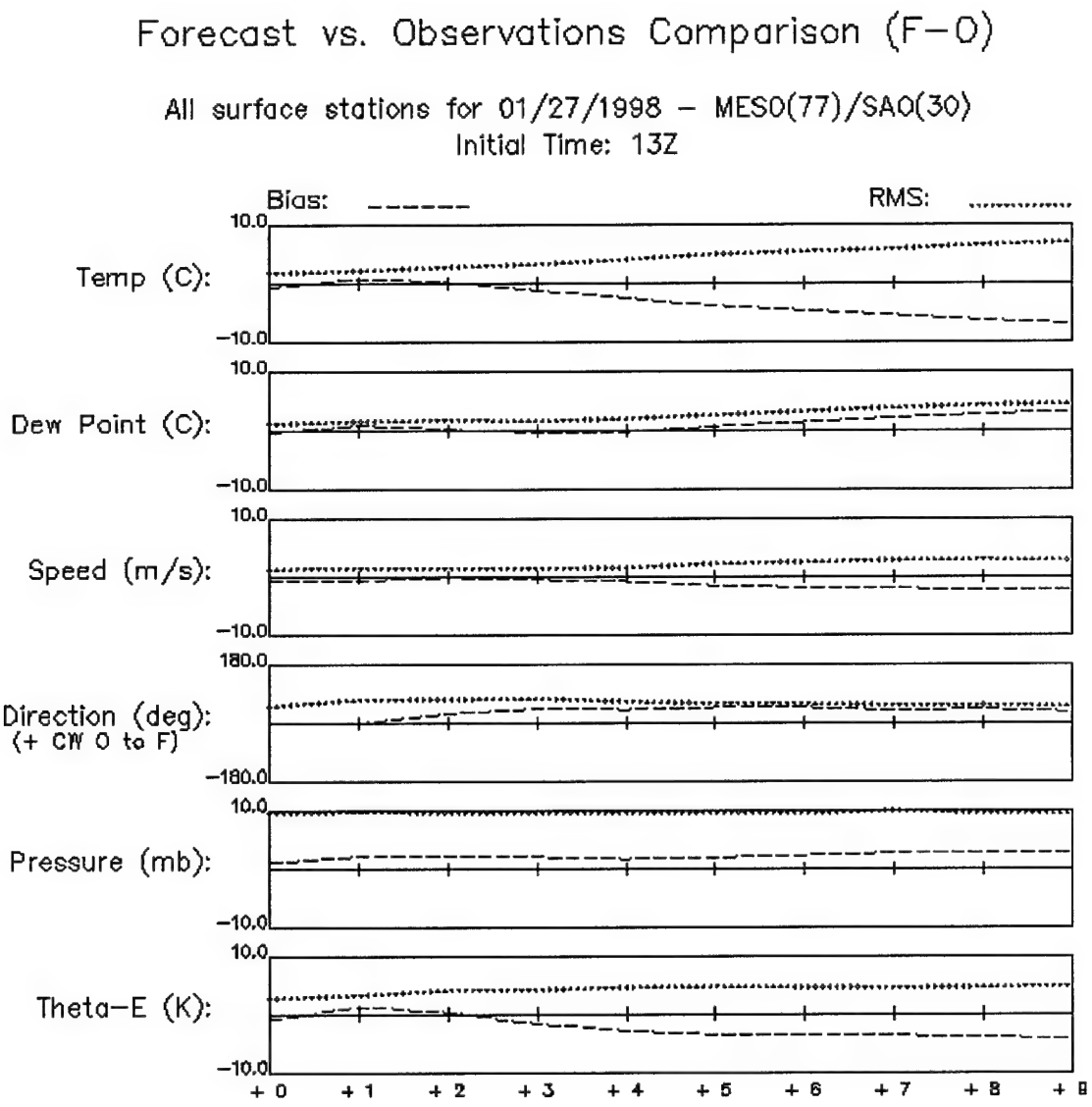


Figure 4.3: Same as Figure 4.1 except only averaged over surface stations for the forecast made on 27 January 1998.

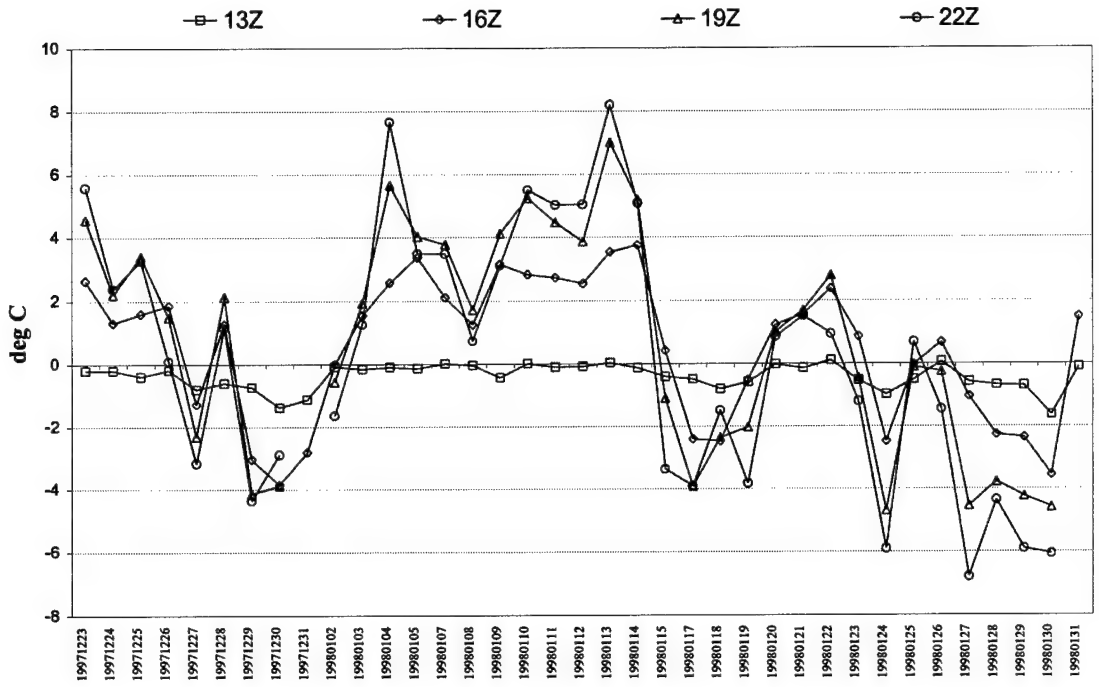


Figure 4.4: Temperature bias (°C) for each forecast day, all surface stations combined, for the 00, 03, 06, and 09 hour forecast points (13, 16, 19, and 22 UTC valid time respectively). (31 December and 31 January forecast only went out to a 03 hour forecast)

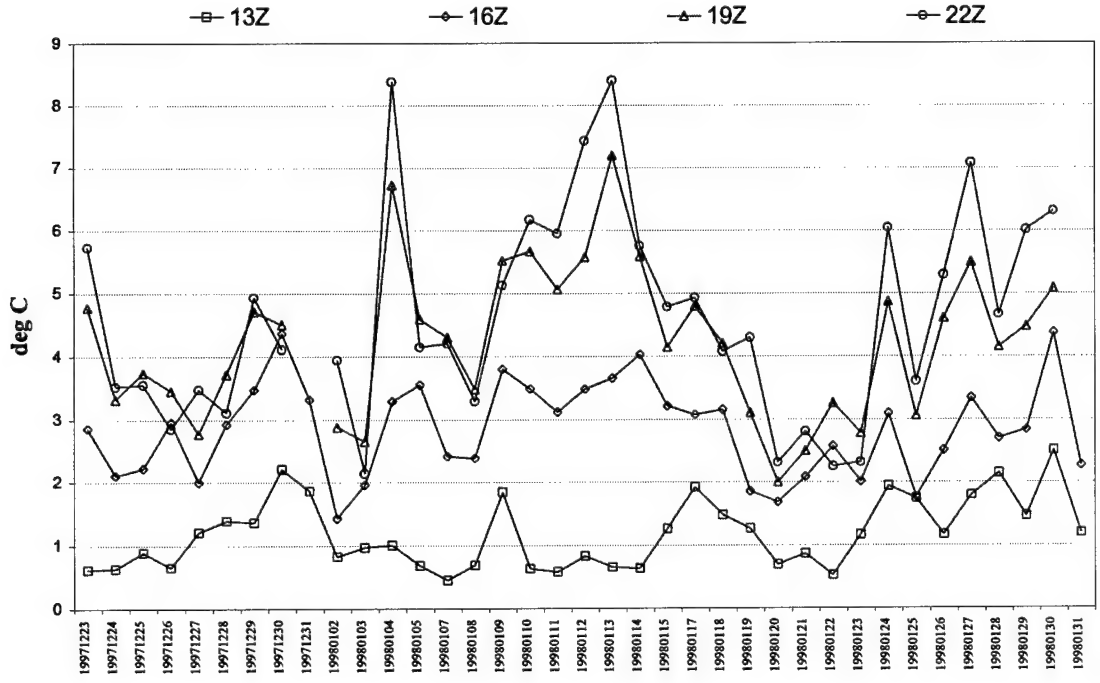


Figure 4.5: Same as Figure 4.4 except for temperature RMSE (°C).

is that on days where the initial analysis field had a low bias, there is almost always a warm temperature bias in the forecast. Conversely, on days where the initial bias starts out negative, a cold bias results. An examination of the weather experienced on those days did not appear to correspond directly to either type of forecast outcome. However, regional cloud cover could have played a role, especially since the majority of stations evaluated (82) were in Oklahoma, and could skew the results based on cloud-cover in the northeast corner of the forecast domain.

Few similar or other unexpected trends were evident in the verification of the parameters other than temperature, shown in Figures 4.6 through 4.15. The moist bias shown in the dew point temperatures, the positive wind direction bias, and the generally negative speed biases are all in agreement with known ARPS deficiencies (Mewes 1997). The only other significant feature evident is the drastic change in station pressure RMSE observed after 15 January. Although the basic framework of the model was held consistent throughout the entire WOP97 period, this sudden jump indicates something happened to the forecasts or to a change in the processing of the observational database. Since pressure is not a variable used by the icing algorithms, this issue is left for a later case study examination of individual forecast days surrounding the transition period.

An isolation of the bias and RMSE values at each surface observation location over the entire set of WOP97 forecasts is shown in Figures 4.16 through 4.27. In these figures the abscissa (horizontal axis) does not display all (107) station names due to space limitations. The surface airways stations are on the left side (indicated with the K preceding the standard three letter identifier) and the Mesonet stations are on the right, separated by the vertical line displayed. Within each group the stations are listed in

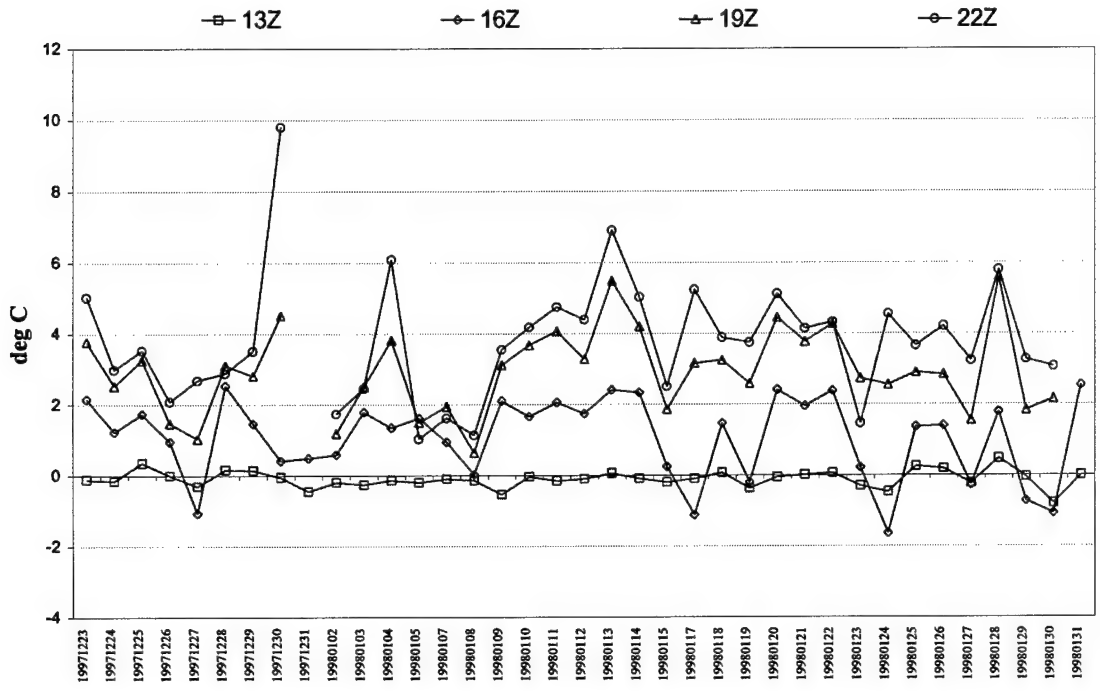


Figure 4.6: Same as Figure 4.4 except for dewpoint Bias (°C)

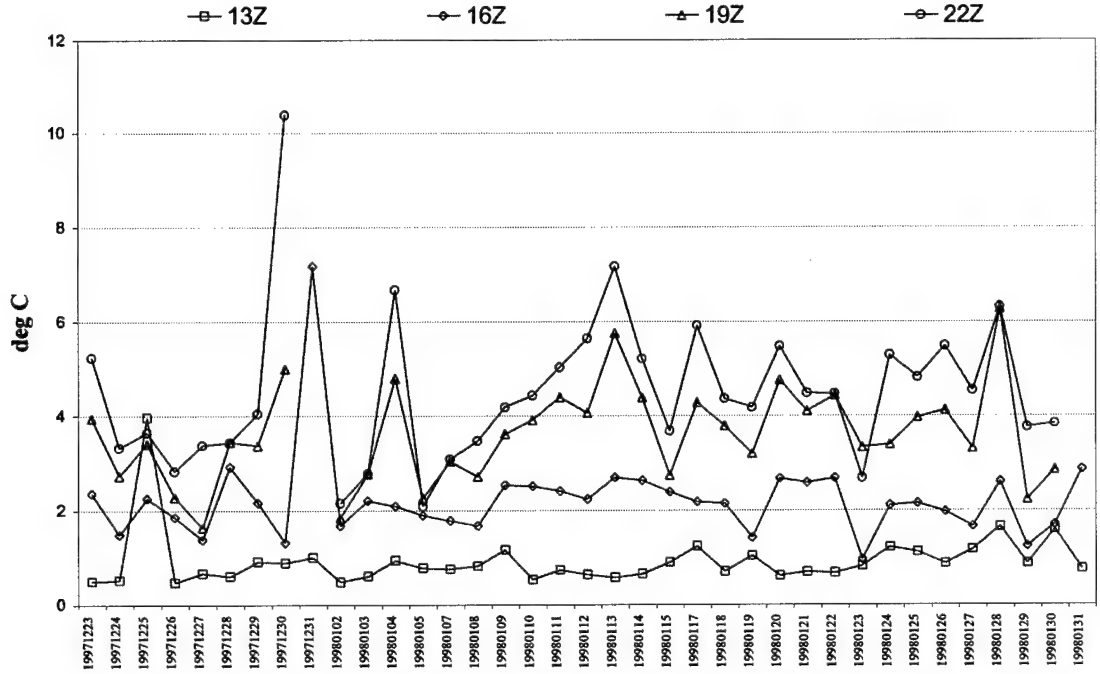


Figure 4.7: Same as Figure 4.4 except for dewpoint RMSE (°C).

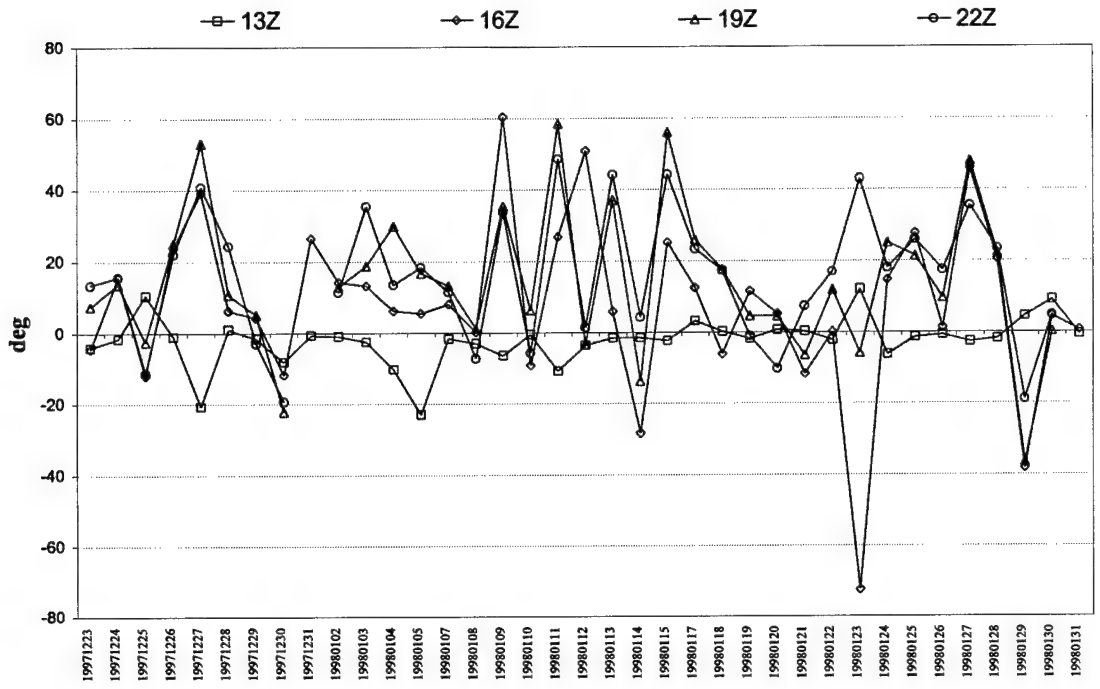


Figure 4.8: Same as Figure 4.4 except for wind direction Bias (°)

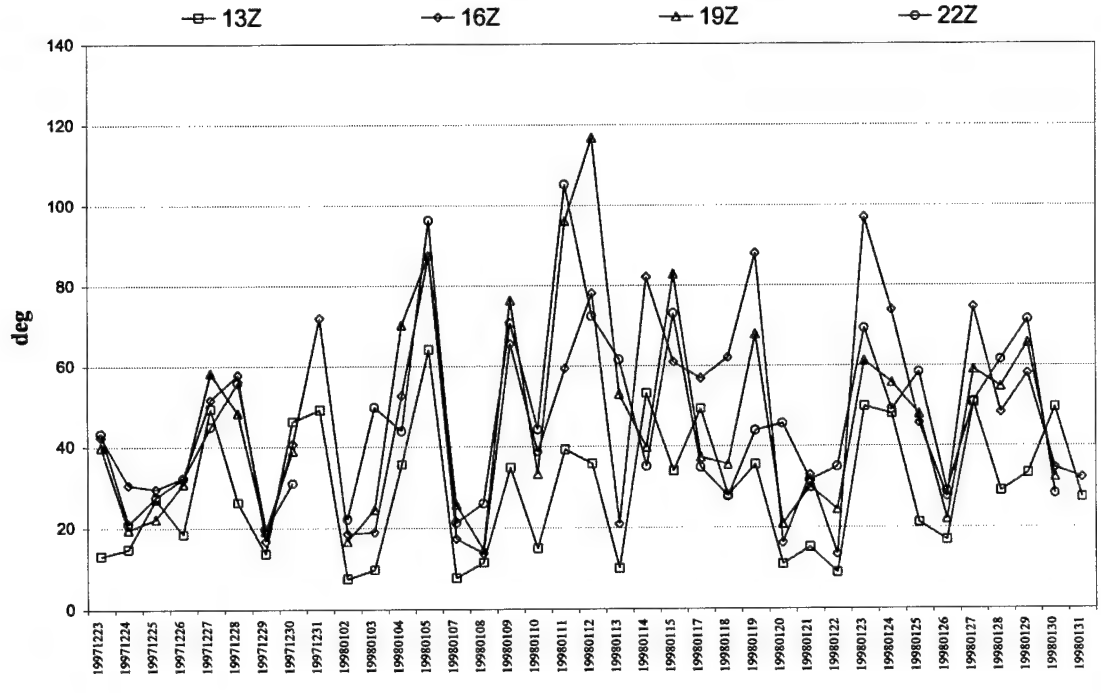


Figure 4.9: Same as Figure 4.4 except for wind direction RMSE (°).

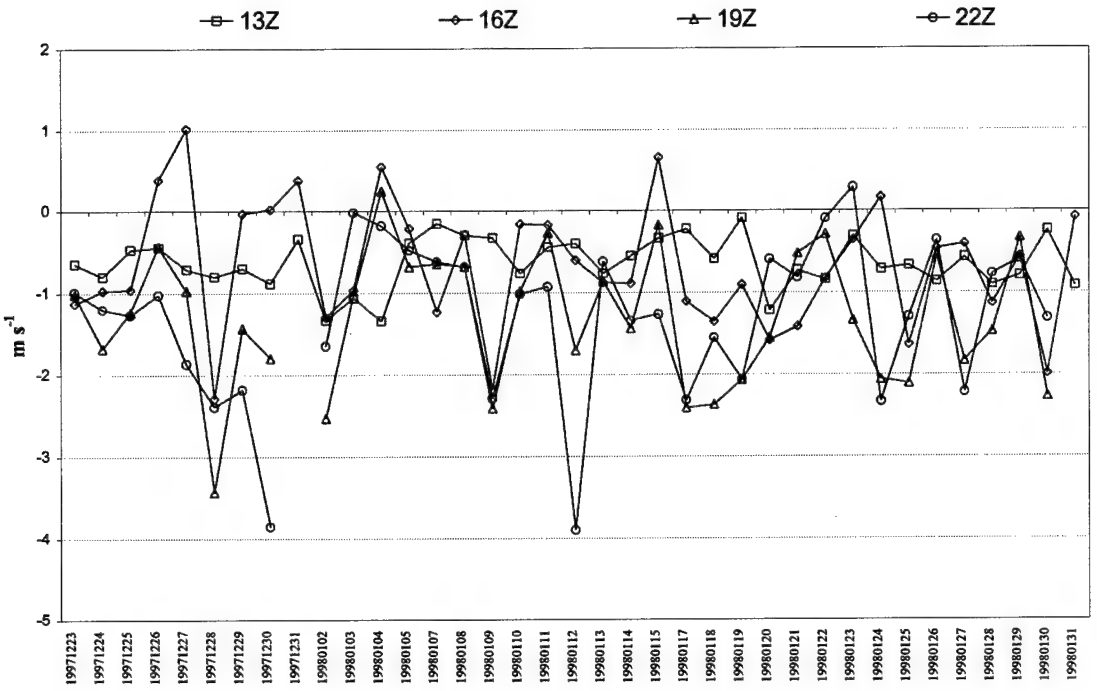


Figure 4.10: Same as Figure 4.4 except for wind speed Bias (m s^{-1})

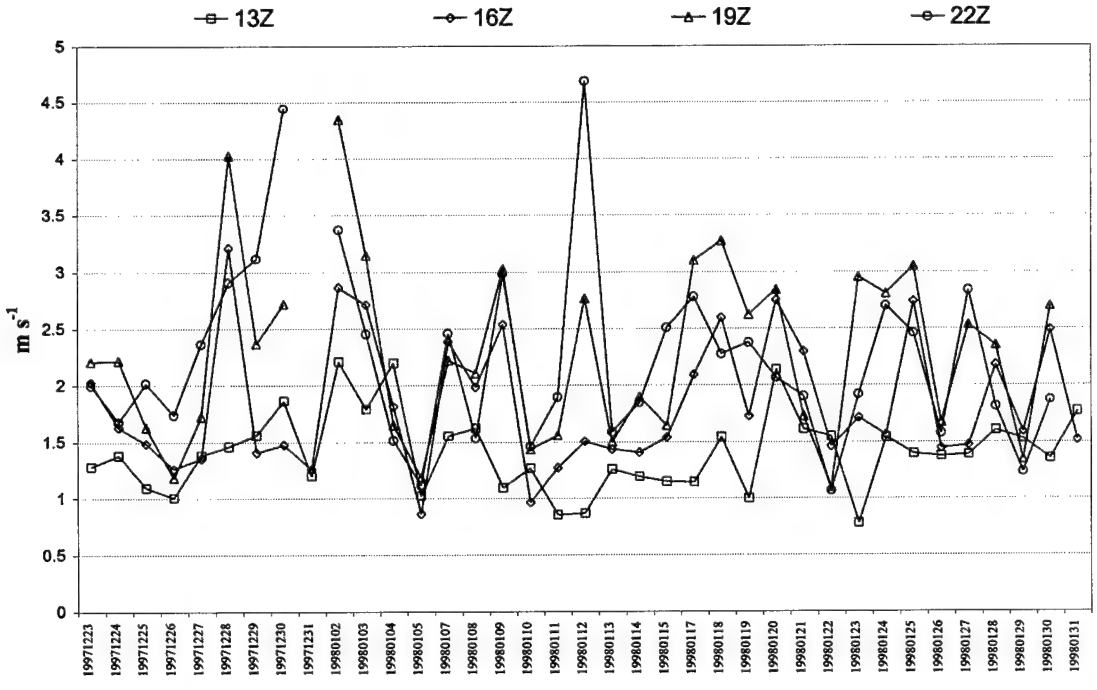


Figure 4.11: Same as Figure 4.4 except for wind speed RMSE (m s^{-1}).

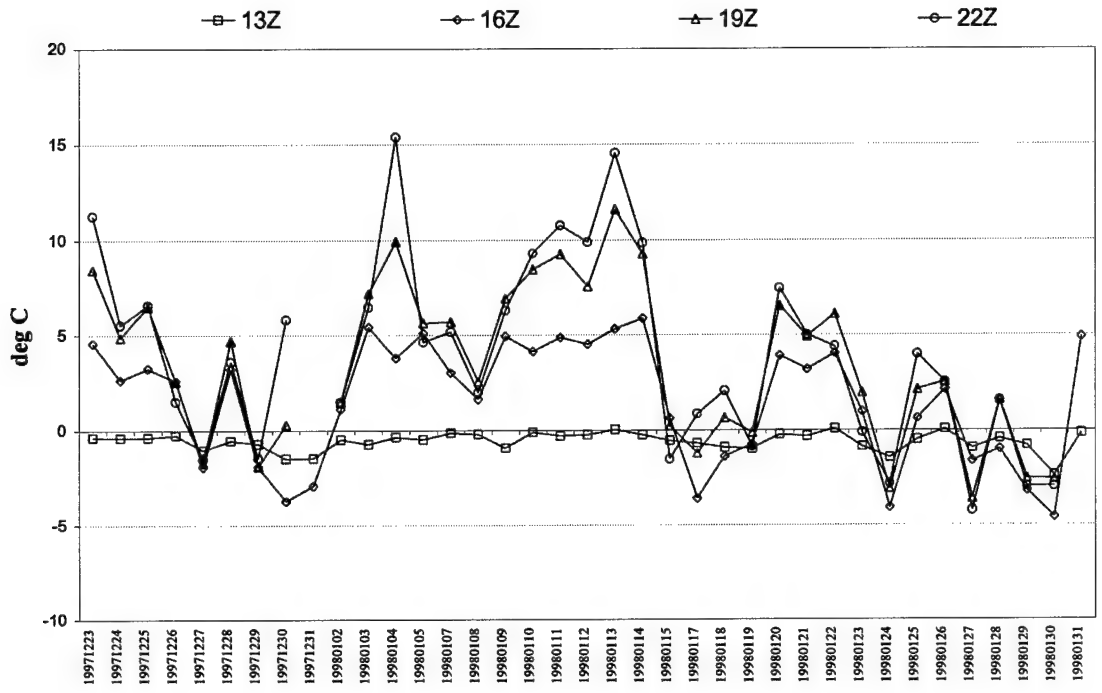


Figure 4.12: Same as Figure 4.4 except for Equivalent Potential Temperature Bias (K)

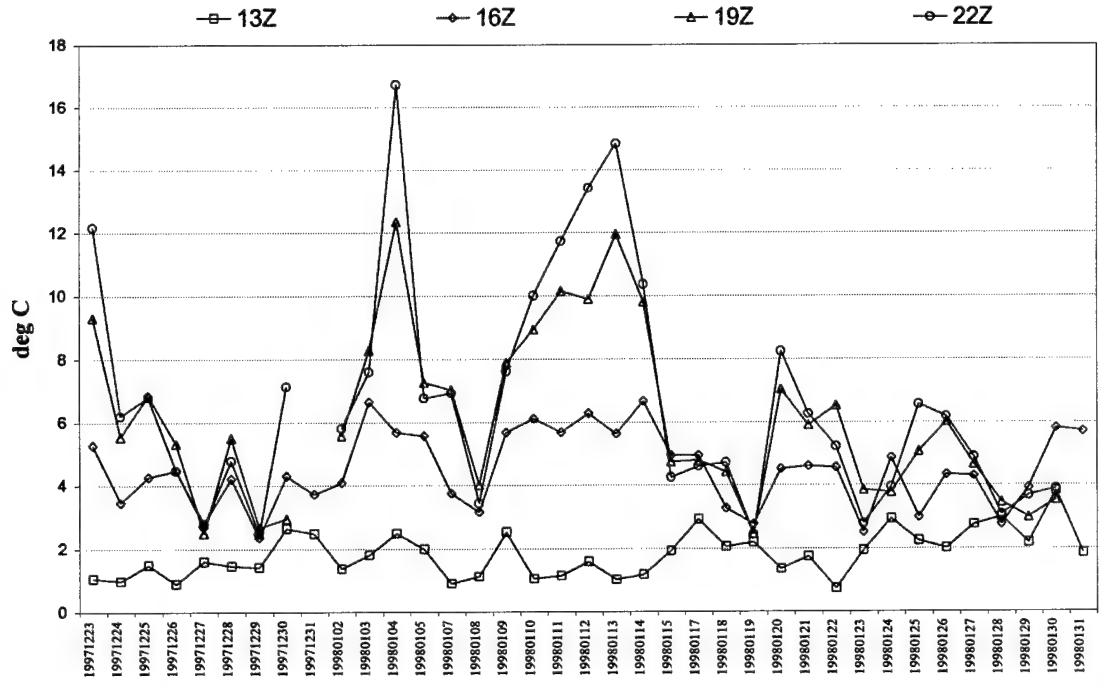


Figure 4.13: Same as Figure 4.4 except for Equivalent Potential Temperature RMSE (K).

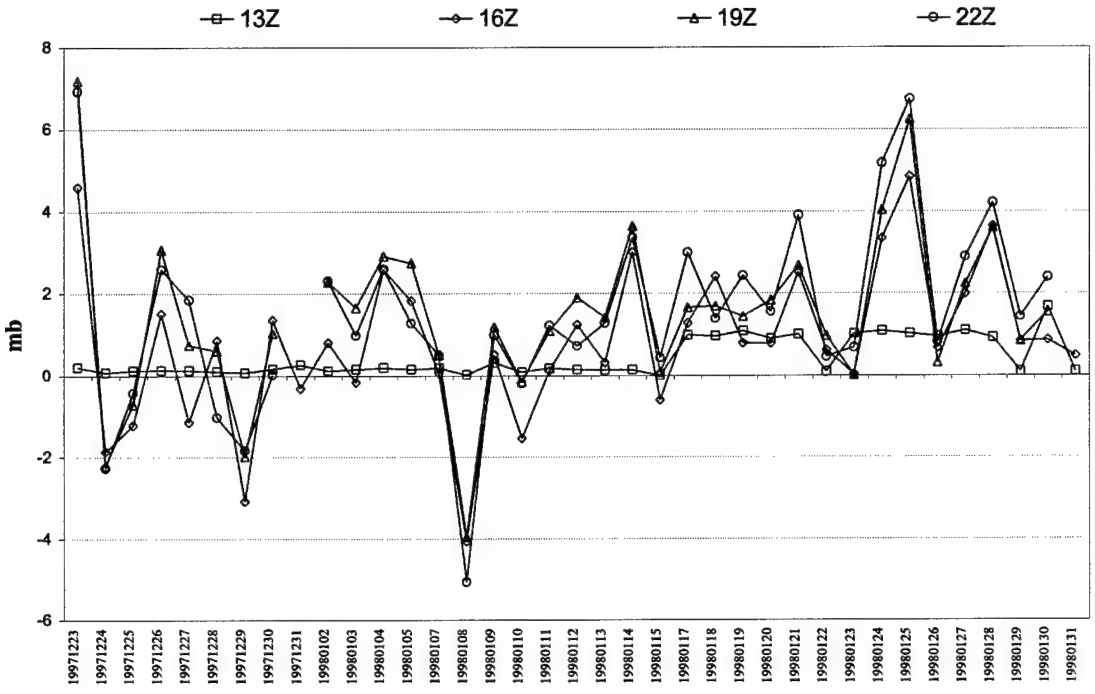


Figure 4.14: Same as Figure 4.4 except for station pressure bias (mb)

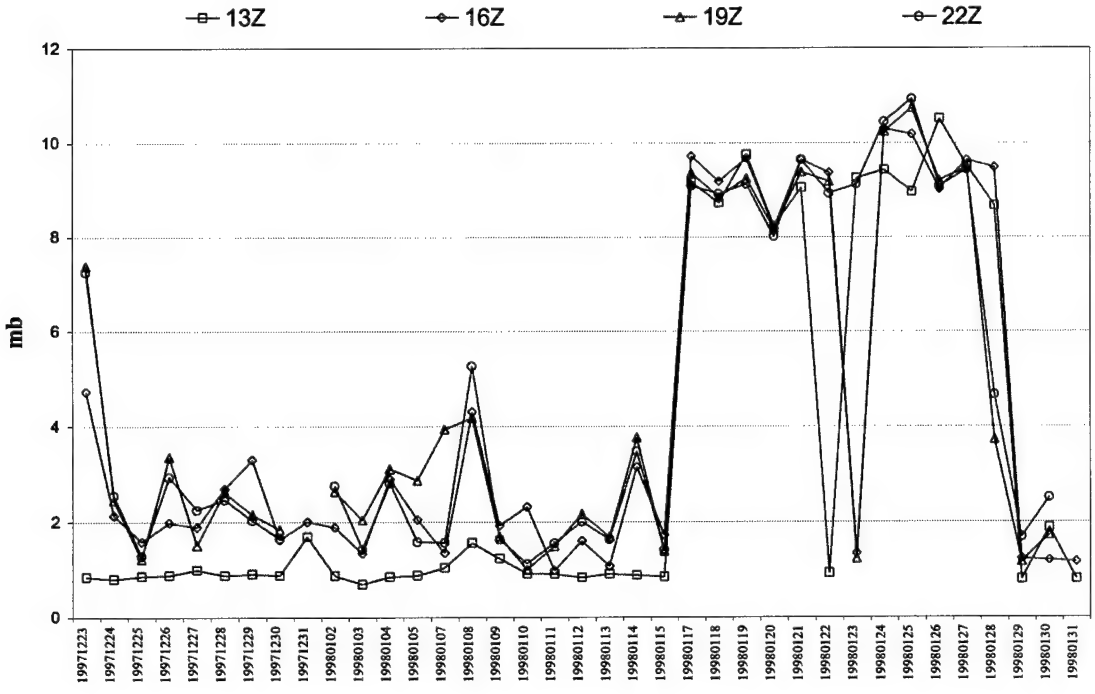


Figure 4.15: Same as Figure 4.4 except for station pressure RMSE (mb).

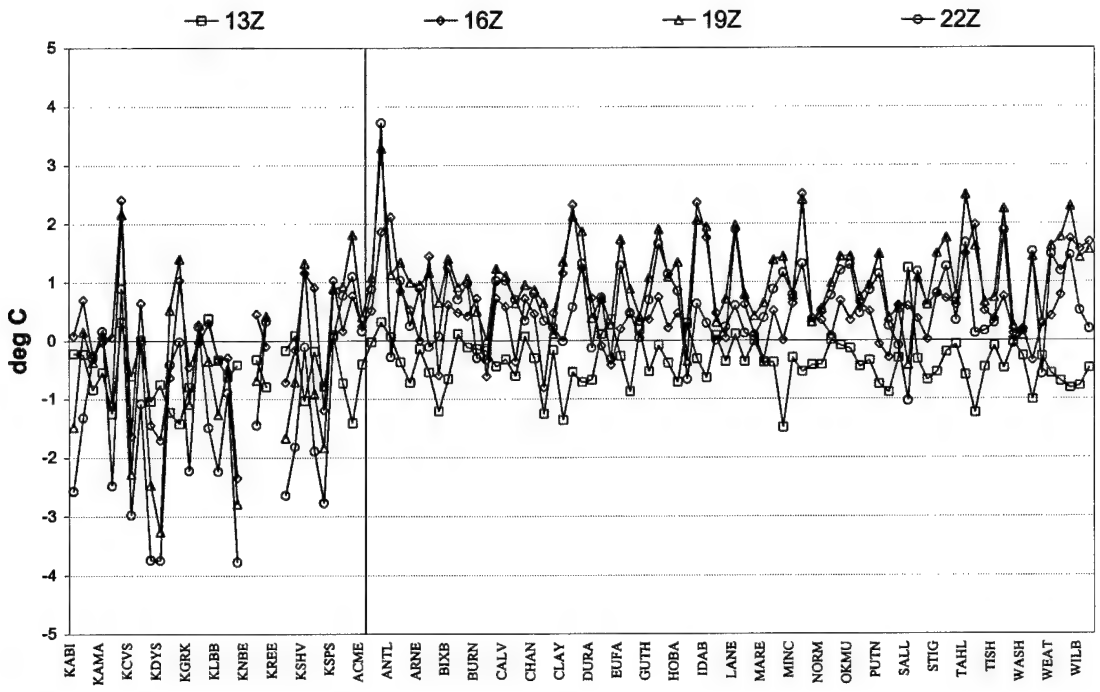


Figure 4.16: Temperature bias (°C) for each surface station (not all identified on axis) for the 00, 03, 06, and 09 hour (13, 16, 19, and 22 UTC respectively) over the entire WOP97.

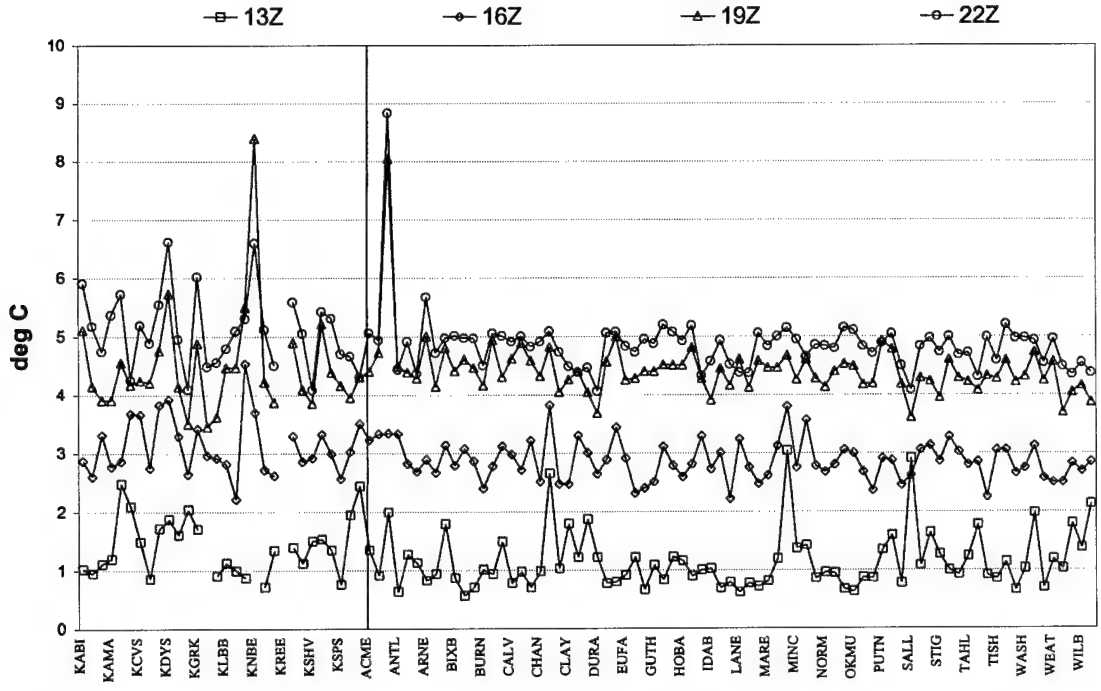


Figure 4.17: Same as Figure 4.16 except for temperature RMSE (°C).

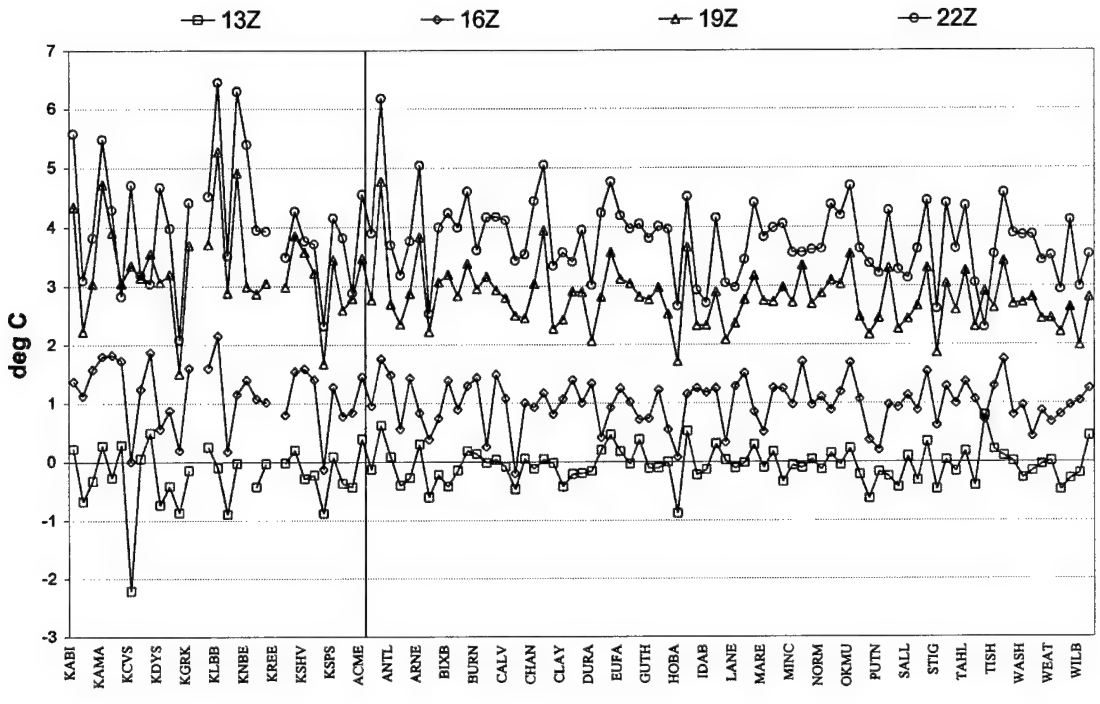


Figure 4.18: Same as Figure 4.16 except for dewpoint bias (°C)

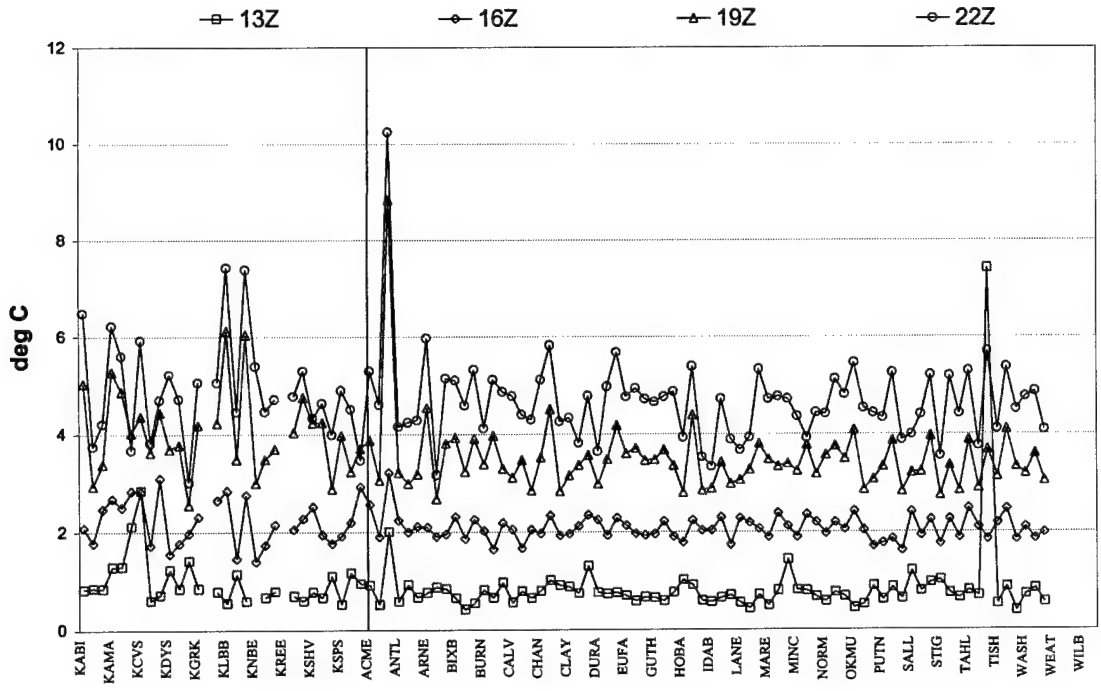


Figure 4.19: Same as Figure 4.17 except for dewpoint RMSE (°C).

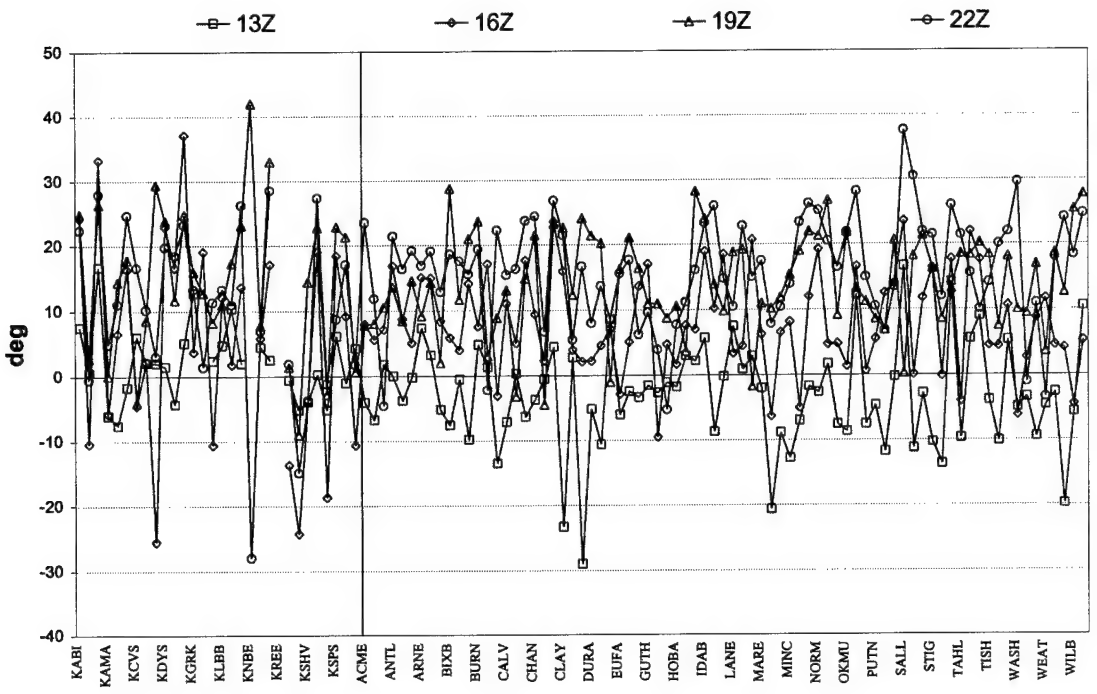


Figure 4.20: Same as Figure 4.16 except for wind direction bias (°)

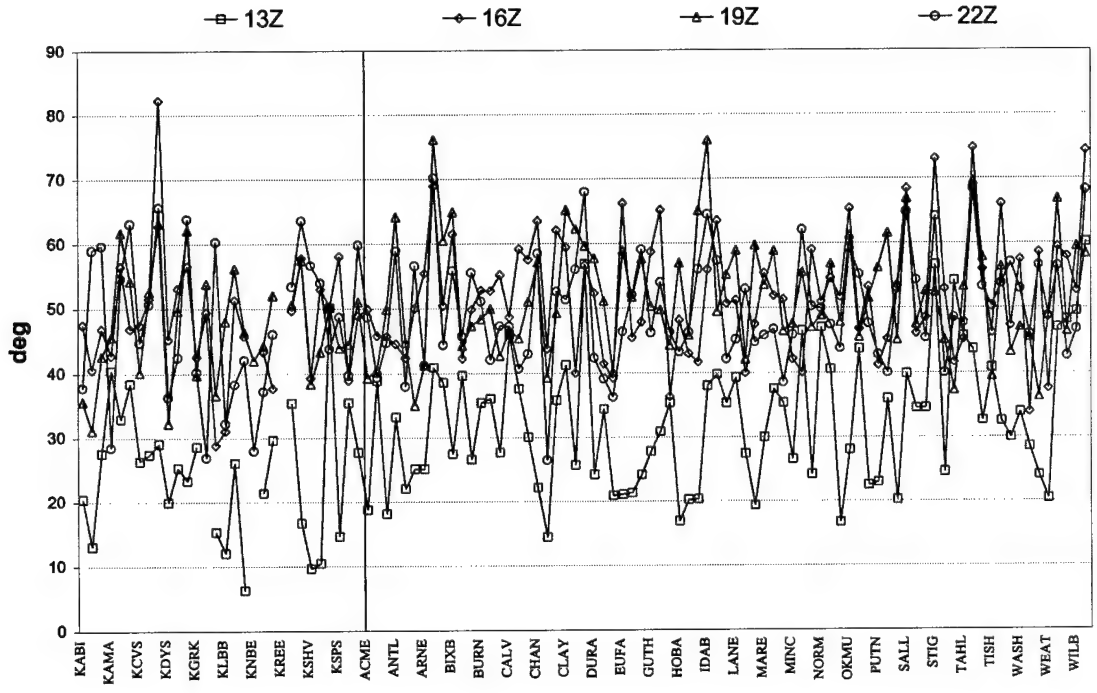


Figure 4.21: Same as Figure 4.17 except for wind direction RMSE (°).

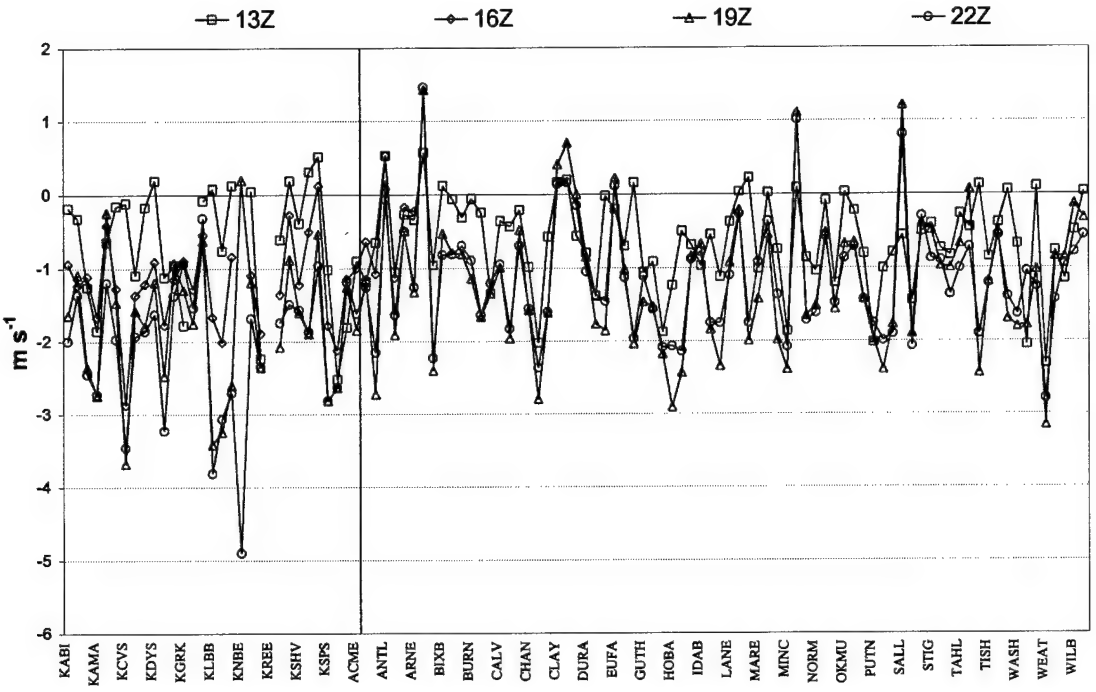


Figure 4.22: Same as Figure 4.16 except for wind speed bias (m s^{-1})

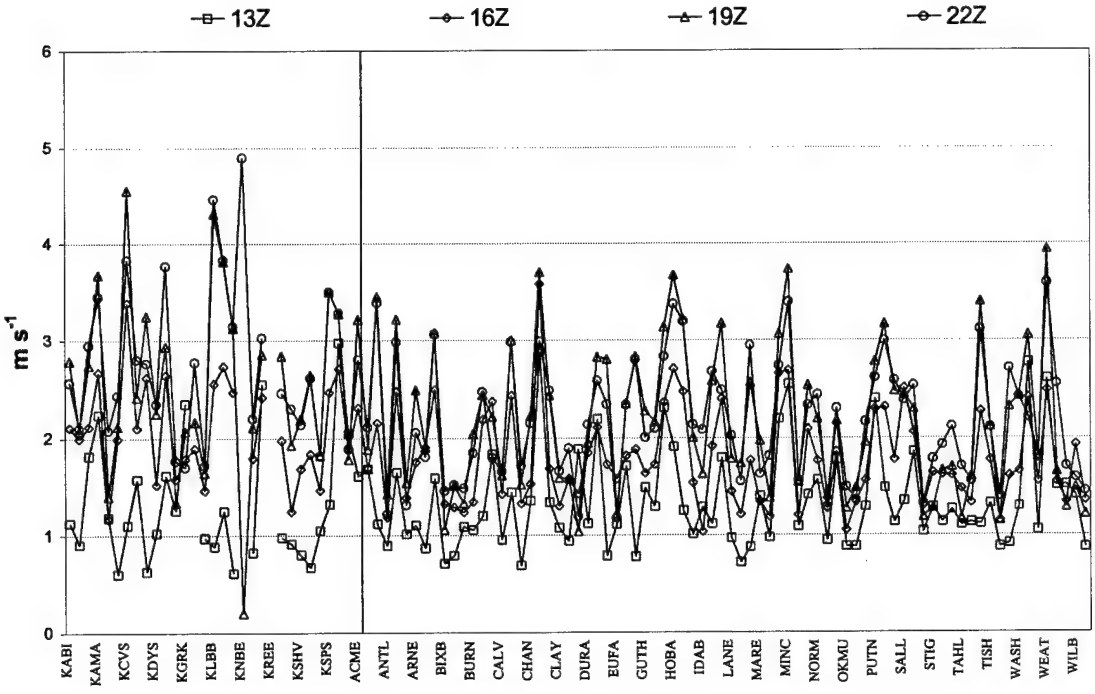


Figure 4.23: Same as Figure 4.17 except for wind speed RMSE (m s^{-1}).

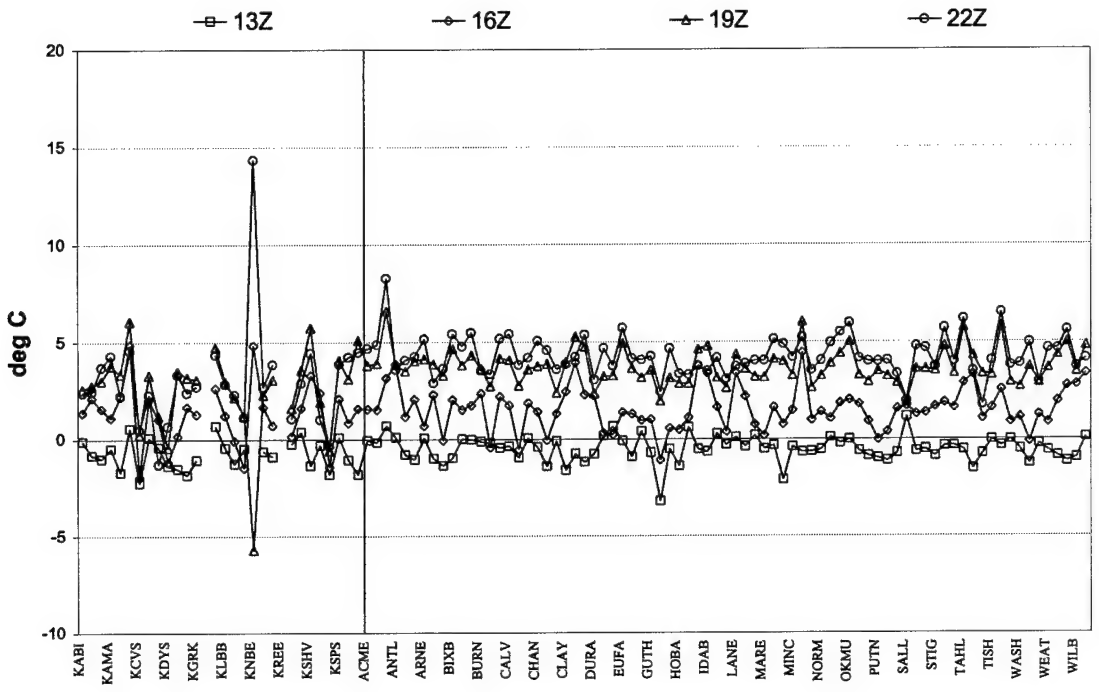


Figure 4.24: Same as Figure 4.16 except for equivalent potential temperature bias (K)

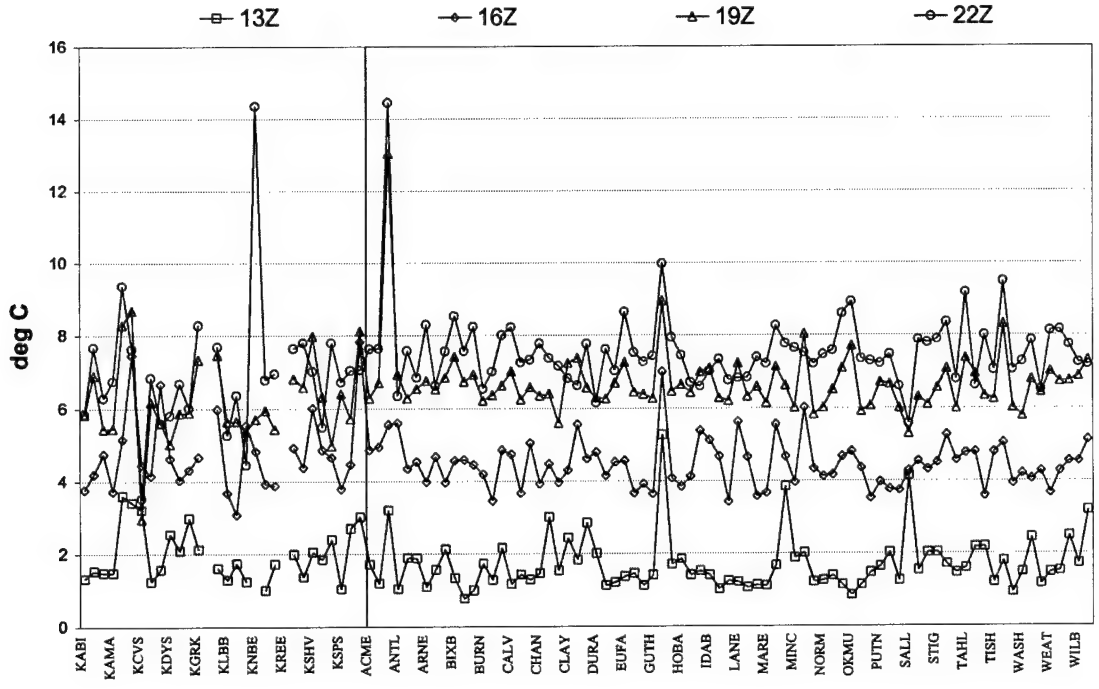


Figure 4.25: Same as Figure 4.17 except for equivalent potential temperature RMSE (K).

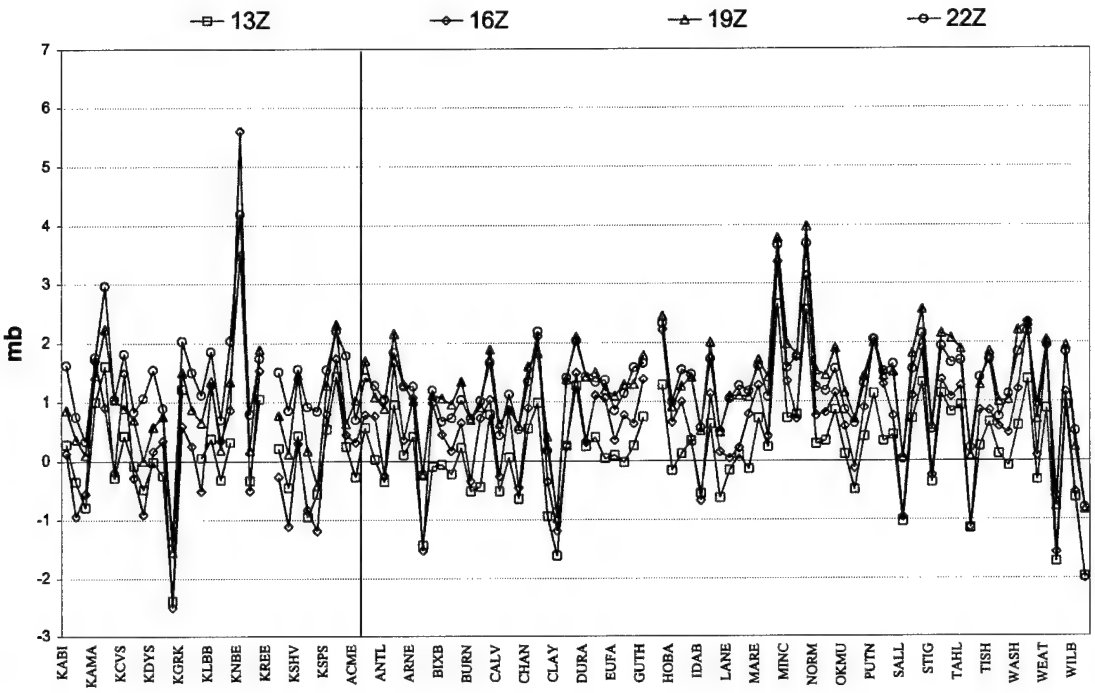


Figure 4.26: Same as Figure 4.16 except for station pressure bias (mb)

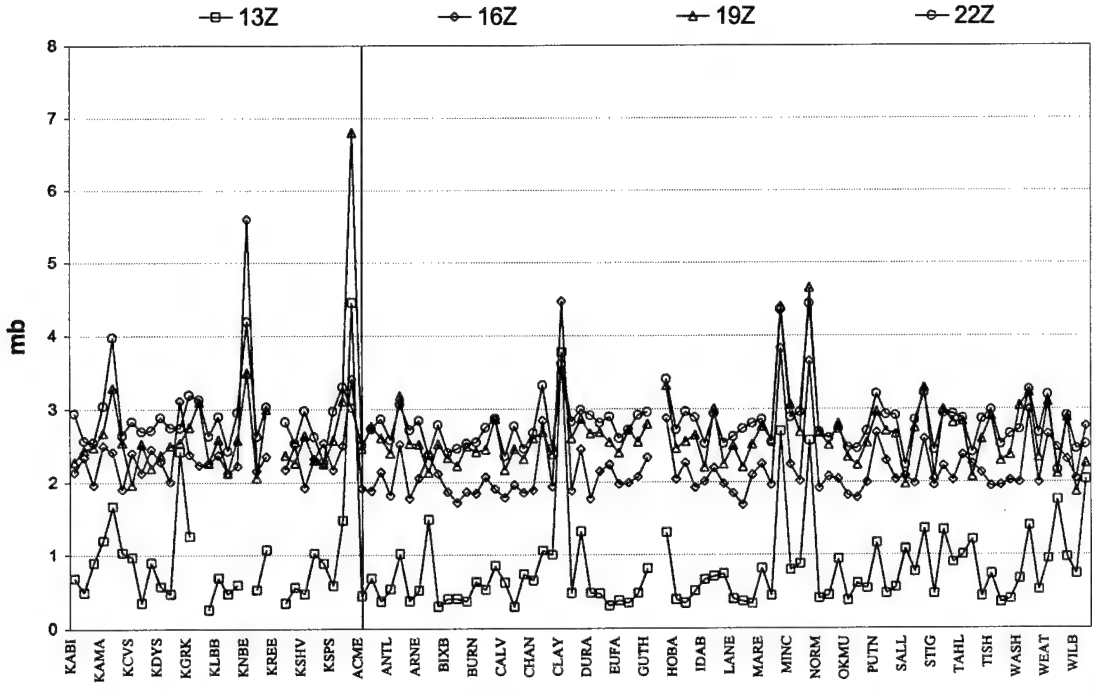


Figure 4.27: Same as Figure 4.17 except for station pressure RMSE (mb).

alphabetical order from left to right. The results show that there is general similarity among all observations located throughout the domain. The trends and characteristics evident in the analysis of daily numbers are still present throughout the domain. The only variation again is in the temperature comparison. Here there is a reversal in bias sign between the Mesonet sites (positive bias) and the surface airways stations (negative bias). One explanation for this could be the differences in reporting precision of the observing equipment. The routine surface stations report integer values of temperature while the Mesonet observations are to the nearest tenth of a degree. If this was a factor we would see the same result in the dew point results in Figure 4.18. However, the dew point doesn't show the same characteristic. An examination of the error meteograms at each station showed that the limited spatial coverage of the Mesonet sites simply results in them having more similarity to the other reports from stations in Oklahoma, than to the remaining sites in central and southern Texas. Figure 4.28 shows one sample from each group (the Mesonet site at Medicine Park; Tinker AFB, OK; and Randolph AFB, Texas) which illustrate this.

As a final look at the surface verification characteristics, contingency plots showing raw forecast/observation pairs of temperature, dew point, and wind speed and direction were created. In this manner, a look at the variability and development trend of forecast errors can be examined across the entire range of values experienced for each variable. Figures 4.29 through 4.32 show these contingency plots, which display data pairs in an (x,y) relationship for three hour intervals. In the plots shown, the x-coordinate is the applicable observed value, while the y coordinate is the value that was forecast. Thus the resulting display is of the form (f_i,o_i). A linear regression, best fit indication is

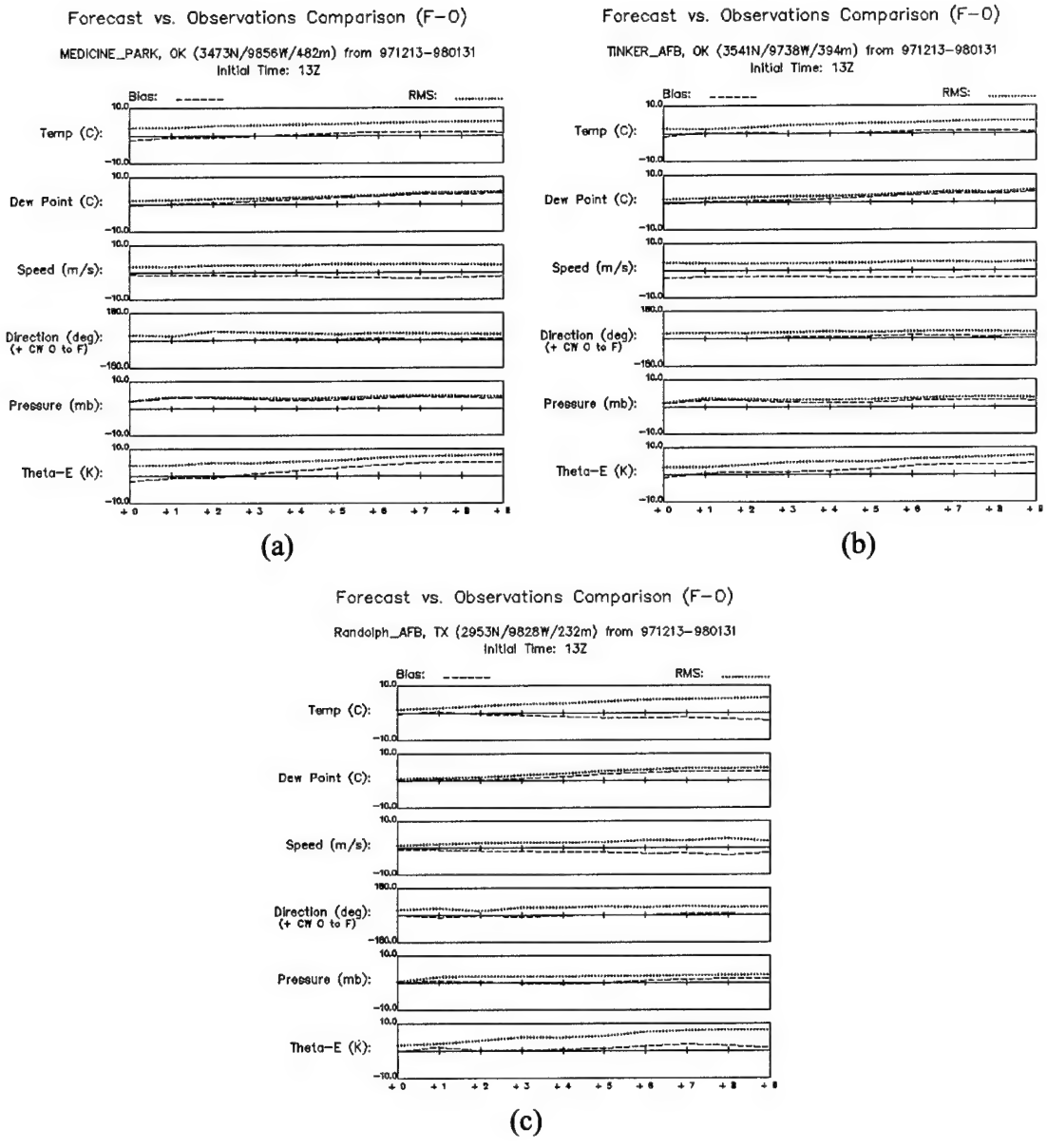


Figure 4.28: Average bias and RMSE for (a) Medicine Park, OK, (b) Tinker AFB, OK, and (c) Randolph AFB, OK, over all WOP97 forecasts combined.

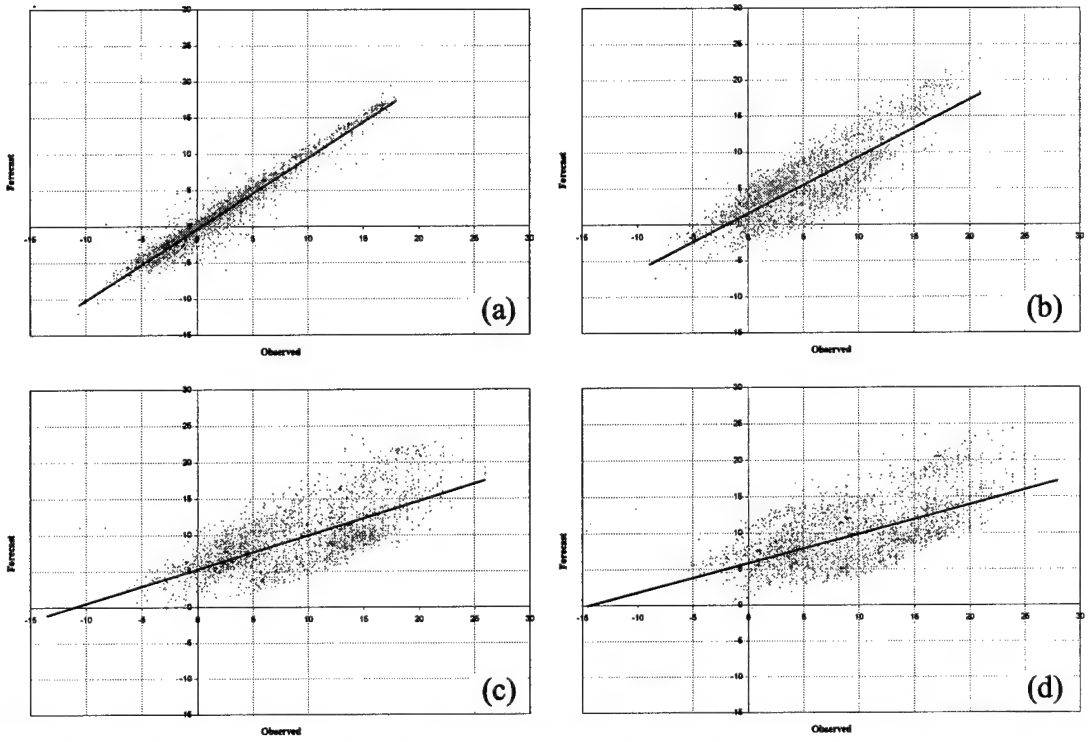


Figure 4.29: Temperature (°C) contingency plot for all surface forecast/observed data pairs (Mesonet and SAO) for the (a) 00, (b) 03, (c) 06, and (d) 09 hour forecasts (13, 16, 19, and 22 UTC respectively) for entire WOP97.

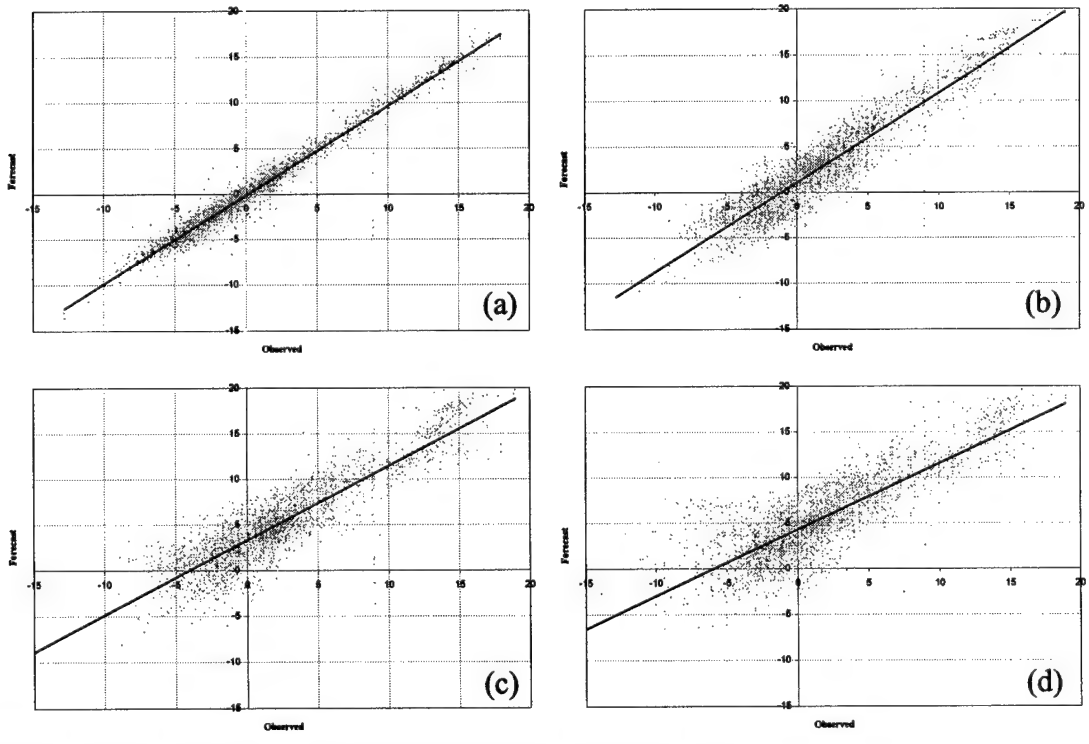


Figure 4.30: Same as Figure 4.29 except for Dewpoint (°C).

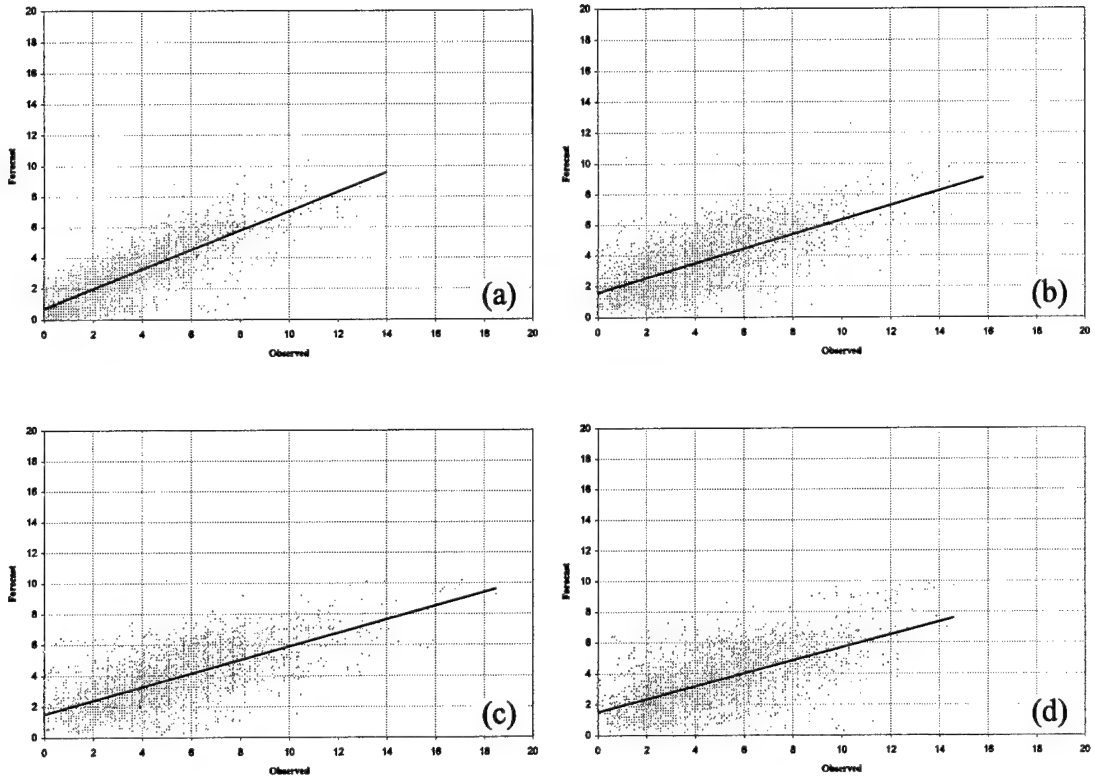


Figure 4.31: Same as Figure 4.29 except for Wind Speed (m s^{-1}).

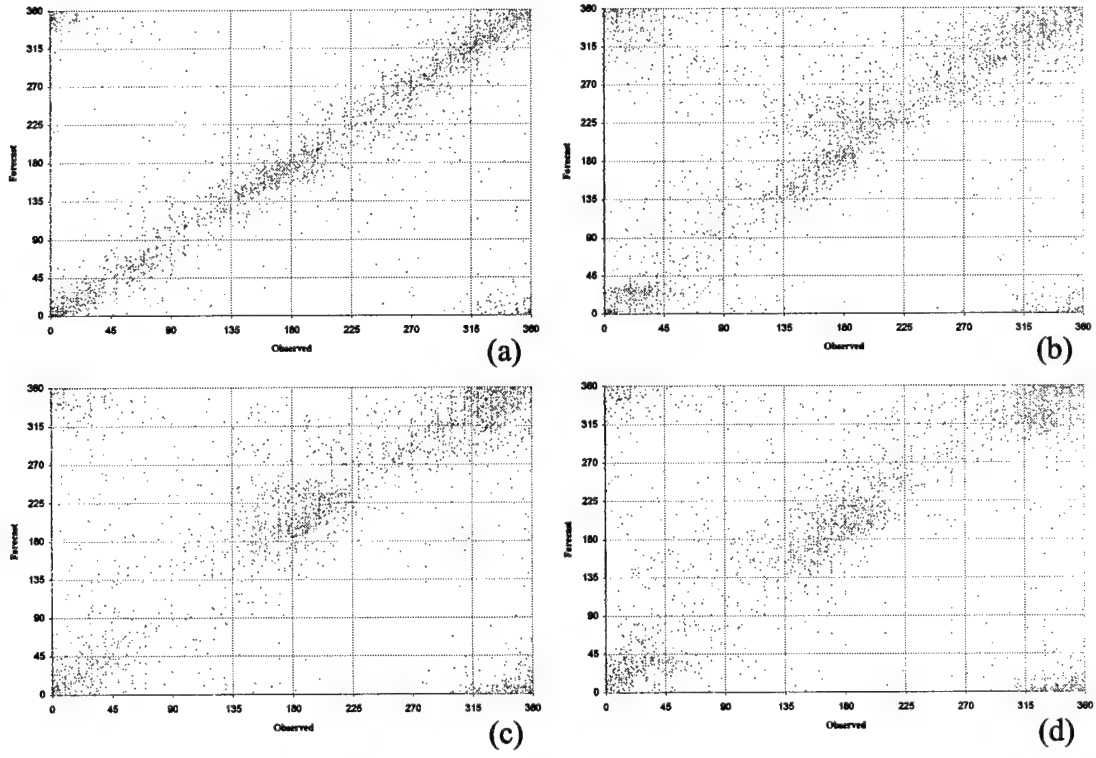


Figure 4.32: Same as Figure 4.29 except for Wind Direction (°).

also shown by the solid black line for all plots except wind direction due to the ambiguity of direction values around 360°. A perfect forecast would be indicated by x = y (i.e. a straight line from lower left to upper right corners respectively) with little or no variability around it. For each variable shown, the initialization data (00 hour forecast) shows a good relationship across the entire range of values. The regression lines generally move upward (i.e. toward top of graph) for temperature and dew point with time, showing an increases in the overall bias as the forecast progresses. For wind speeds the trend is slightly downward indicating the forecast wind speeds are less than the observed values, and get worse with time. These trends match those shown earlier in the daily averages, with the forecast errors increasing with time. The plots do, however, indicate that the error trends are not uniform across the range of observed values. The slopes of the regression lines decrease with time, showing forecasts get warmer (moister) in the lower observed ranges while getting cooler (dryer) in the higher ranges. Thus for temperature, the "warmer" temperature region shows an increasing "cold" bias with time, while for "cooler" temperatures the trend is to increase the "warm" bias. The opposite nature of these two regions may account for the small average magnitude of the temperature bias in Figure 4.1. The dew point plots show that the "moist" bias seen earlier is more consistent over the entire dew point range. The "moist" bias is especially evident in the lower dew point range. The variability about the regression line increases for all parameters as the forecast progresses as we would expect. For the wind direction, while the basic positive bias is present across all directions, there is more variability in forecast error when the wind is observed to be from the north, with more consistent errors when a southerly direction is observed.

Section 4.2 Verification of Upper-Air Forecasts

(a) *Wind Profiles*

The second verification data set used in this work, the wind profiler network observations, provides the first (and regretfully, the only) upper air comparison possible with the WOP97 forecasts. Although there is no direct input of the horizontal winds into any of the icing algorithms, an indication of accurate winds may imply that the thermodynamic profiles that are used in the icing assessments can also be considered "accurate". Thus, for each reporting height in the vertical, the bias, RMSE, MVE, and the RMVSE, along with the "relative" vector errors, MRVE and RMSRVE, were computed to examine ARPS wind forecasts.

The comparison of the wind speed and direction (same error convention as surface wind direction) averaged for the entire set of WOP97 forecasts is given in Figures 4.33 through 4.36. The figures have the height (in meters) along the abscissa, with the corresponding verification measures plotted on the ordinate. Here we see that the overall quality of the forecasts are good, with the magnitude of the bias values below 1 m s^{-1}, and RMSE values uniform, and relatively low as well. These match earlier results of other model comparisons to wind profiler observations (Richardson 1993). The magnitude of the RMSVE errors, 10-15m s^{-1}, are however, much larger than the 4-8 m s^{-1} seen in the earlier comparisons. The short duration of the operational period and the smaller resolution of the ARPS forecasts may account for some of these differences.

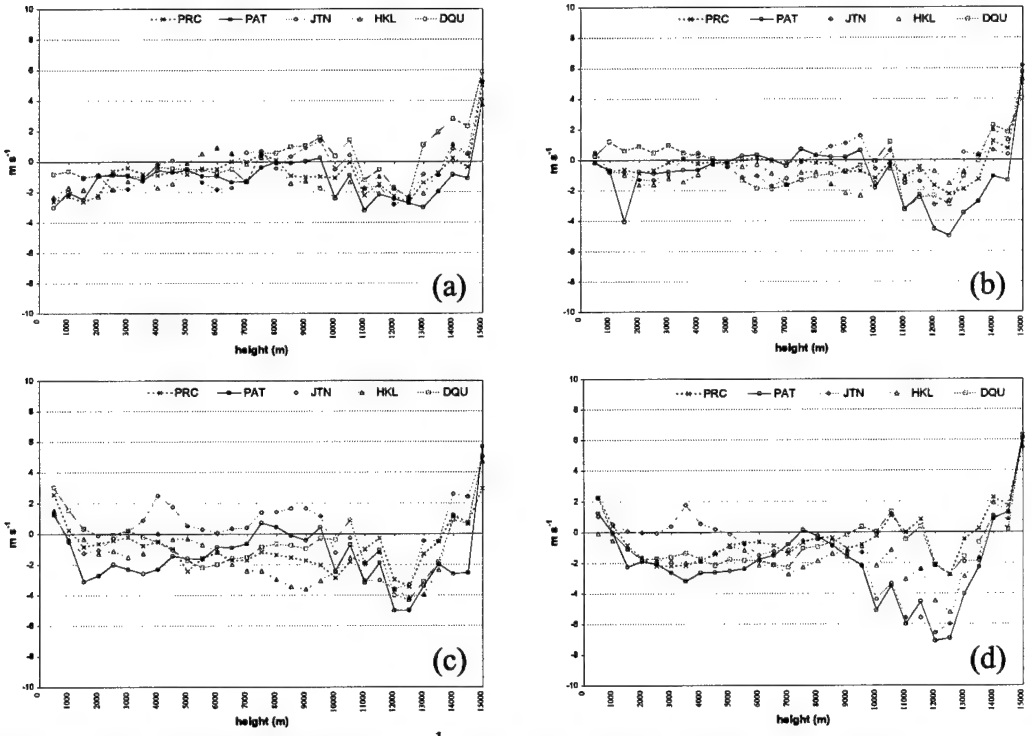


Figure 4.33: Wind speed bias (m s^{-1}) for each profiler location for (a) 00, (b) 03, (c) 06, and (d) 09 hour (13, 16, 19, and 22 UTC respectively) forecast over all days combined.

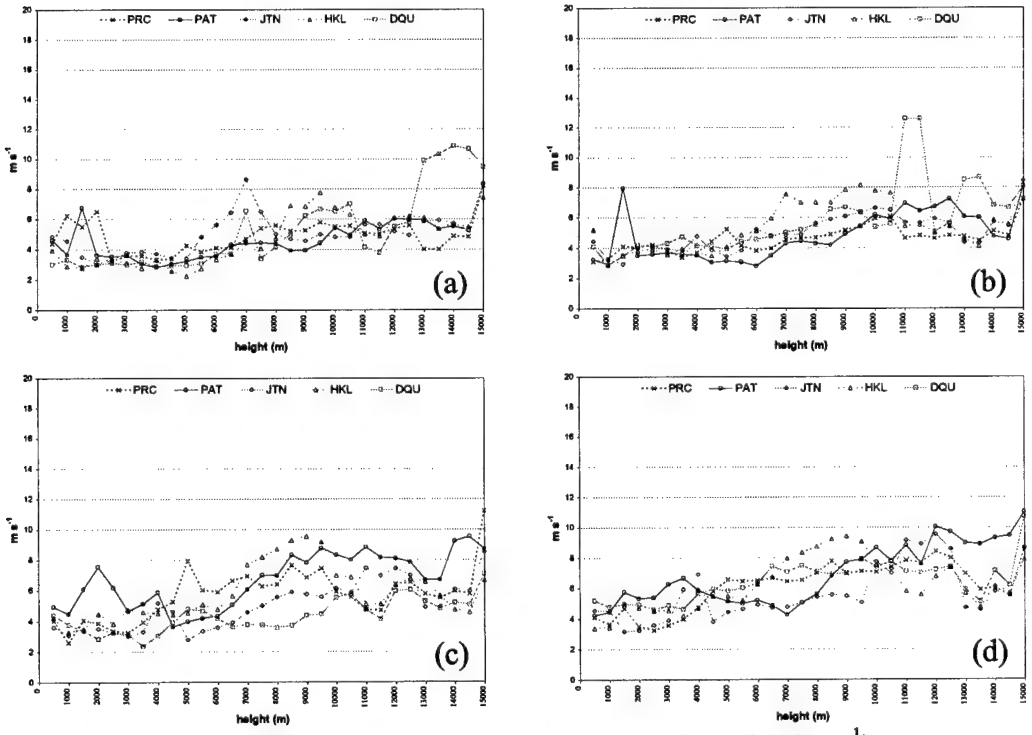


Figure 4.34: Same as Figure 4.33 except for wind speed RMSE (m s^{-1})

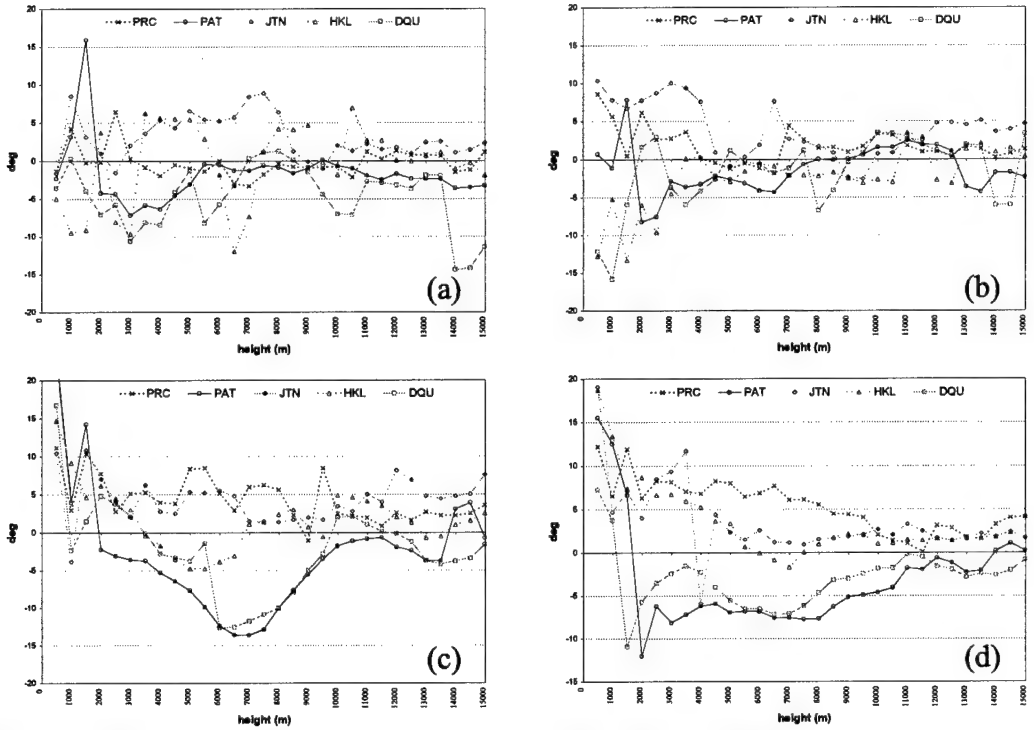


Figure 4.35: Same as Figure 4.33 except for wind direction bias (°)

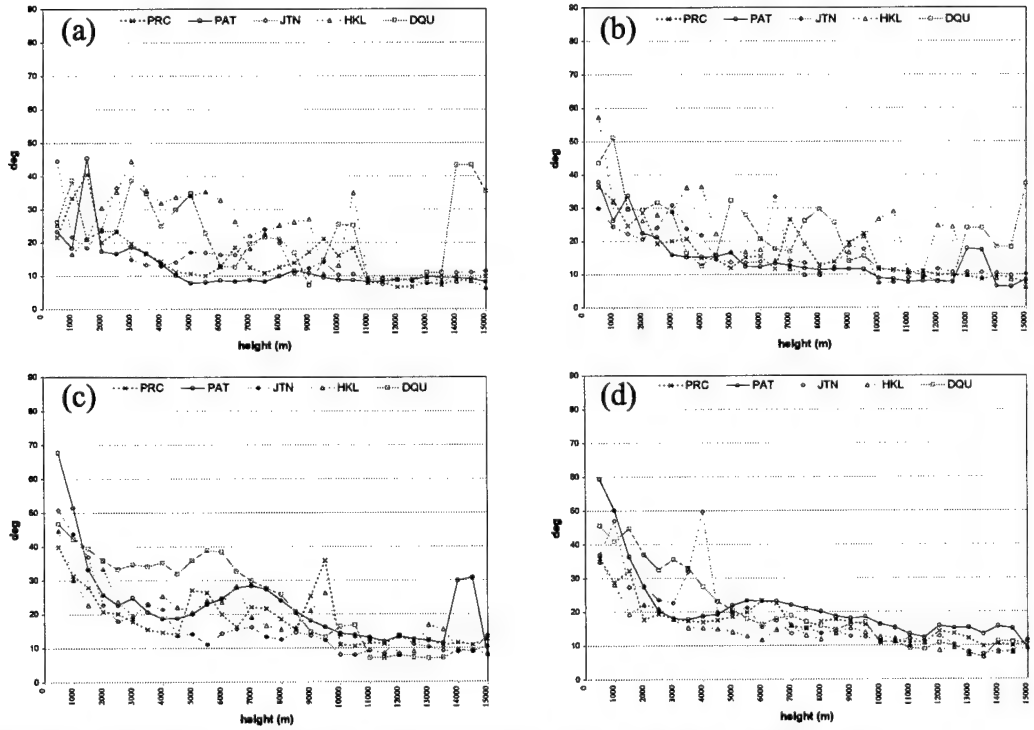


Figure 4.36: Same as Figure 4.33 except for wind direction RMSE (°)

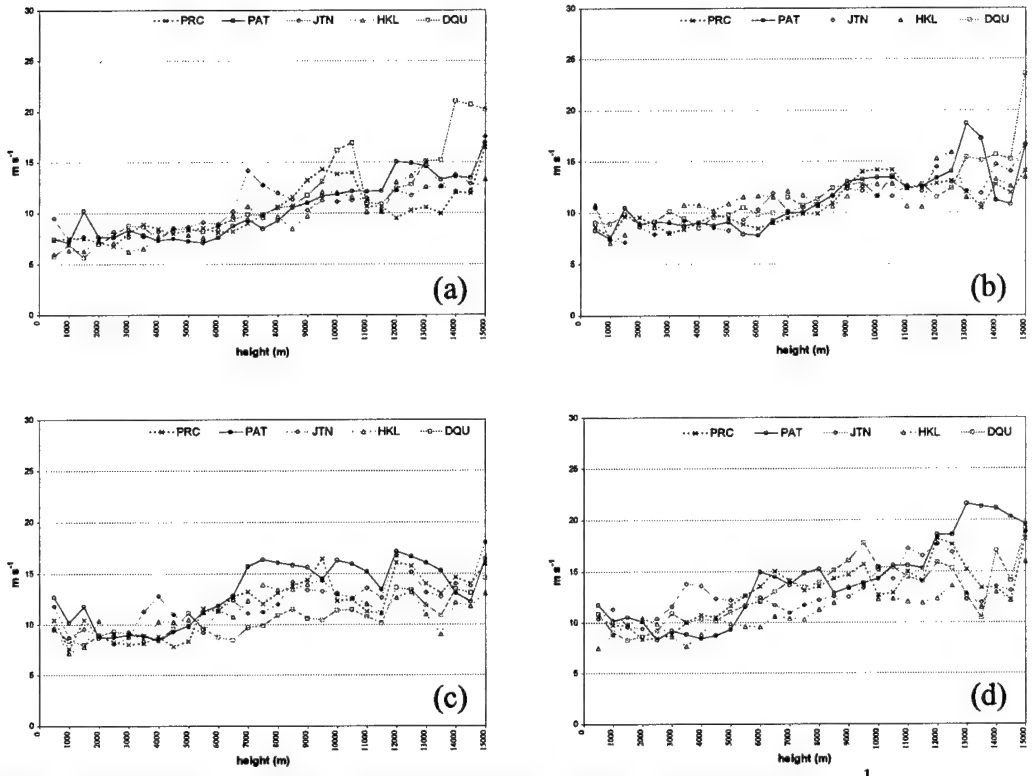


Figure 4.37: Same as Figure 4.33 except for average wind MVE (m s^{-1})

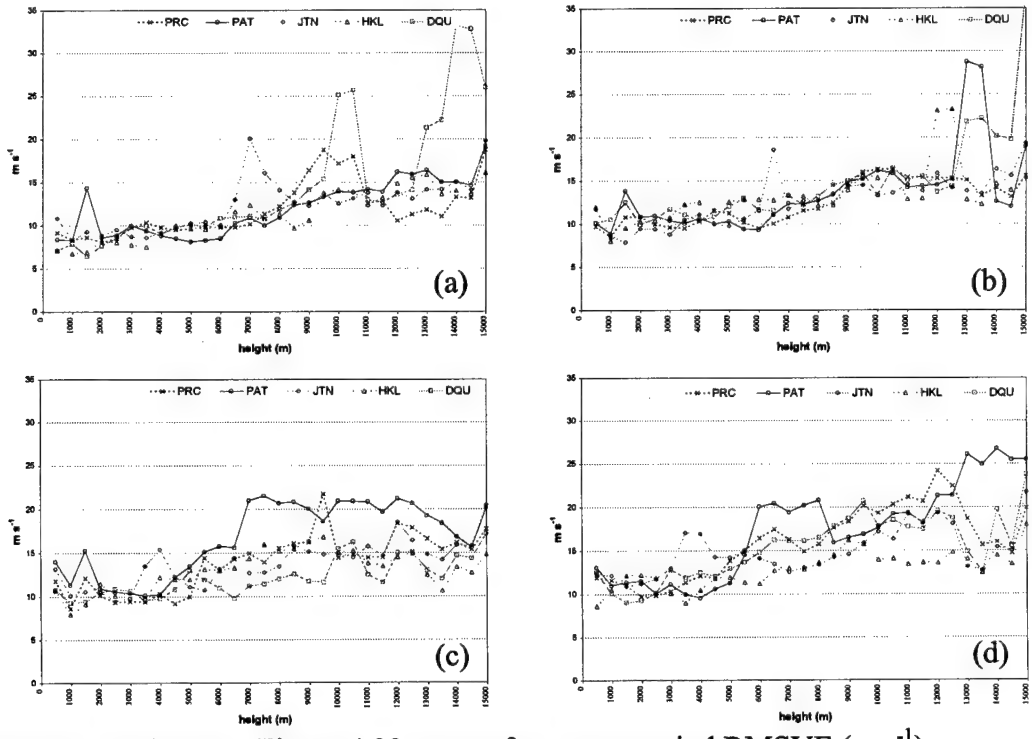


Figure 4.38: Same as Figure 4.33 except for average wind RMSVE (m s^{-1}).

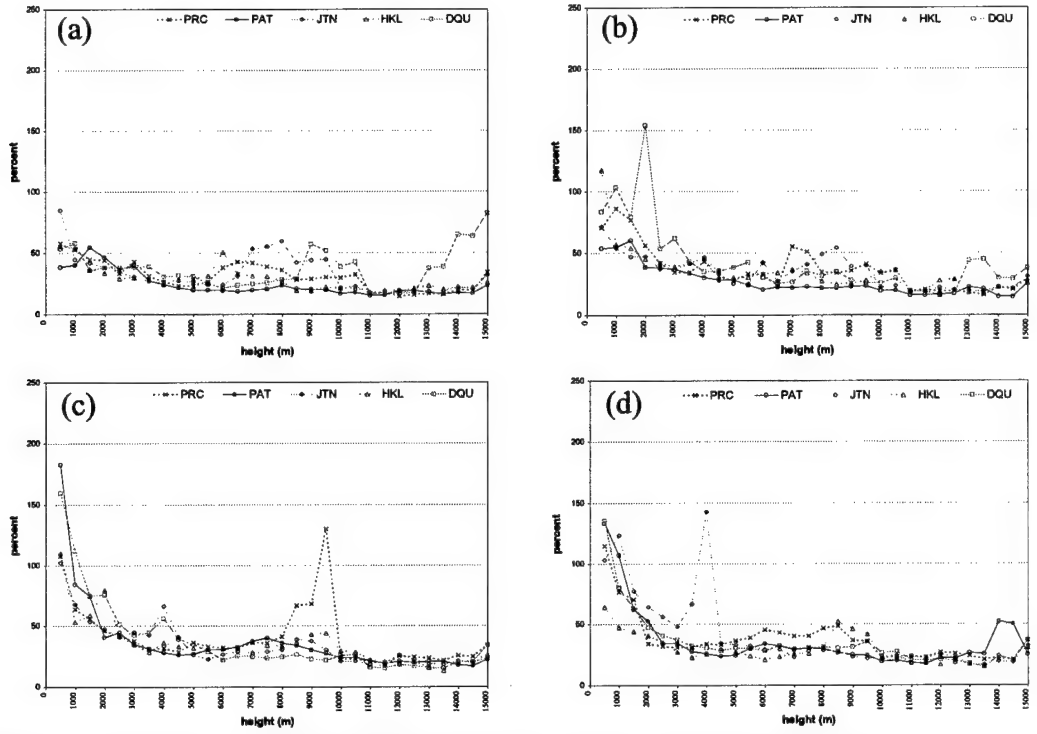


Figure 4.39: Same as Figure 4.33 except for average wind MRVE (%)

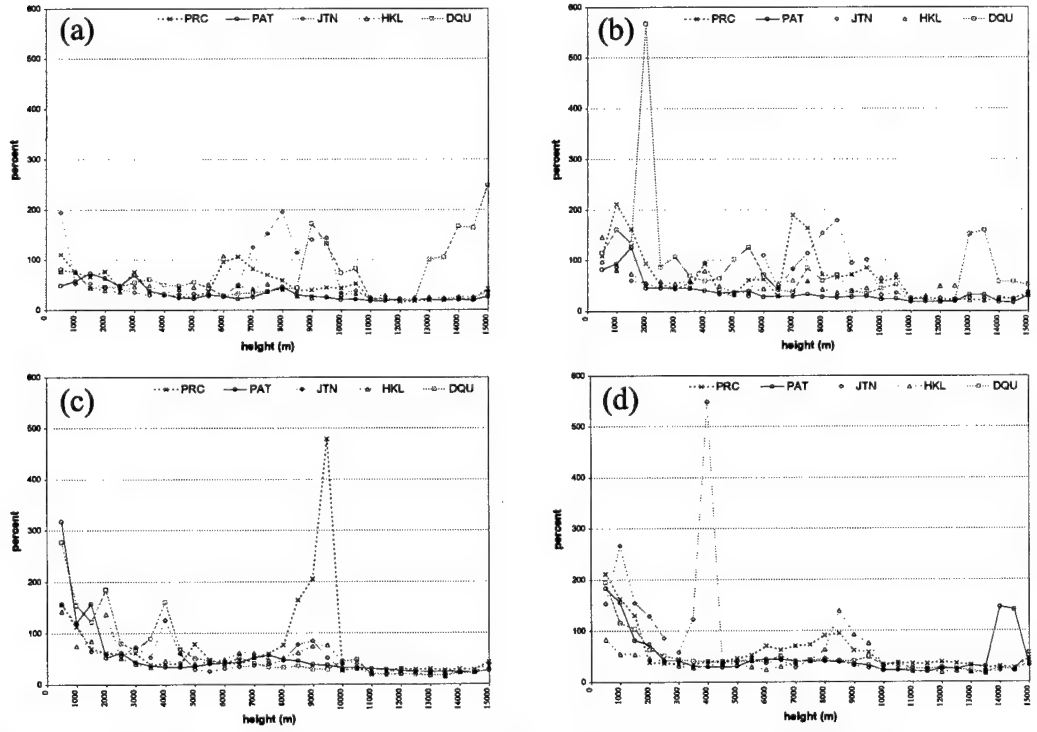


Figure 4.40: Same as Figure 4.33 except for average wind RMSRVE (%)

The speed errors increase slightly with height. For wind speed, the increase in magnitudes with height is plausible since the actual wind speeds themselves increase in height. The opposite is true with wind direction where the forecast error is larger near the surface. This may be due to the greater difficulty a model has in predicting boundary layer wind profiles or where wind speeds are slower in general. Calculations of the absolute *vector* measures, Figures 4.37 and 4.38, show the same overall relationship just described, while the relative ones, in Figures 4.39 and 4.40, give a more direct visible representation of the fact that the forecast errors in the low levels are more significant in nature. Here the relative error measures show the greatest variability and largest forecast error in the low levels of the model, where errors are larger in relation to the actual wind magnitudes. Again, the relative errors gradually decrease with height due to the wind speeds increasing more rapidly with height than the associated errors. We can reason that the same boundary layer forecasting concerns discussed as a cause of the surface forecast errors can explain the why the more significant wind errors are again limited to the lowest levels in the model.

(b) *RUC Analysis Fields*

The final verification made to assess the quality of the WOP97 forecasts is a comparison of ARPS forecasts of variables on constant pressure levels versus RUC analysis data. Of specific interest is the ability of the model to adequately predict the thermodynamic and moisture profiles within the atmosphere. As described earlier, the RUC analyses are the only source of upper air "observations" available during the

Level		Units	007200 (15UTC) BIAS	007200 (15UTC) RMSE	018000 (18UTC) BIAS	018000 (18UTC) RMSE	028800 (21UTC) BIAS	028800 (21UTC) RMSE
850mb	Height	10m	-0.10	1.76	-0.49	2.07	0.52	2.36
700mb			-0.64	2.00	-0.86	2.36	0.01	2.54
500mb			-0.92	2.38	-0.93	2.64	0.03	2.83
300mb			-1.15	2.95	-0.75	3.21	0.48	3.68
Surface	Temperature	°C	-1.79	4.10	-0.09	3.52	-0.60	4.19
850mb			-2.00	2.92	-1.73	2.71	-1.96	3.09
700mb			-0.95	1.41	-0.66	1.31	-0.74	1.41
500mb			-0.23	0.88	-0.05	1.06	0.14	1.19
300mb			-0.37	1.16	-0.08	1.20	0.02	1.35
Surface	Dewpoint	°C	1.48	2.78	2.38	3.36	2.90	4.05
850mb			1.18	4.07	1.24	3.59	0.89	3.38
700mb			2.36	5.18	2.38	4.86	1.92	4.42
500mb			1.83	3.79	1.70	3.65	1.44	3.61
300mb			-0.29	1.06	-0.02	1.12	0.09	1.24
Surface	U Wind Speed	ms^{-1}	0.27	1.69	0.01	2.04	-0.41	2.23
850mb			0.71	2.71	0.79	3.07	0.38	3.24
700mb			-0.70	2.69	-0.61	2.98	-0.76	3.38
500mb			-0.38	2.67	-0.20	3.25	-0.40	3.85
300mb			-0.74	4.23	-1.28	5.31	-2.28	6.71
Surface	V Wind Speed	ms^{-1}	-0.10	1.97	-0.81	2.45	-0.60	2.41
850mb			-0.04	2.73	-1.02	3.37	-1.47	3.54
700mb			0.52	2.62	0.36	2.83	0.09	3.41
500mb			0.55	2.80	0.69	3.50	0.39	4.22
300mb			-0.16	3.65	-0.18	4.59	-0.01	5.55
Surface	Relative Humidity	%	13.52	21.65	8.13	17.08	10.51	20.08
850mb			12.95	20.73	11.65	19.01	11.17	18.84
700mb			11.27	19.63	10.35	19.00	9.02	18.15
500mb			8.61	18.01	7.32	16.90	5.56	16.98
300mb			8.57	17.97	6.40	18.00	1.86	17.33
Surface	SLP	mb	0.69	2.33	0.02	2.24	1.38	2.86

Table 4.1: WOP97 average field bias and RMSE values for surface, 850 mb, 700 mb, 500 mb, and 300 mb pressure levels for each desired quantity at the 02, 05, and 08 hour forecast points (15, 18, and 21 UTC respectively)

WOP97 forecast period. Ideally, a comparison with rawinsonde sounding reports would be the most desirable verification to accomplish.

The computations of bias and RMSE values for the surface and four constant pressure levels for which data were extracted (850, 700, 500, and 300 mb) for: geopotential height, sea level pressure, temperature, dew point temperature, u and v wind components, and relative humidity are summarized in Table 4.1. The results for the surface temperature and dew points are generally consistent with the verification performed on the surface observations. Table 4.1 shows that the magnitude of the temperature bias is generally small. This was evident in the trend displayed in Figure 4.1 earlier. The value of the 02 hour (15 UTC) surface temperature bias of –1.79 °C does appear on to be on the high side, compared to the surface observation results, which was significantly smaller.

For dew points we again see a pronounced surface moist bias that increases significantly with time. A feature which is a bit unexpected is that significant dew point errors are still evident up to the 500 mb pressure level (roughly 18 thousand feet), with the moisture bias at or near 2 °C for the entire forecast. In addition the RMSE values, in many cases more than double the biases, show extensive variability in both the temperature and dew point errors at all levels examined. This does not bode well for the icing algorithms which rely heavily on the profiles of these two quantities to predict icing conditions. There is one saving grace: the comparison of relative humidity (the direct moisture input to many of the algorithms) appears to be better (bias below 15%) than that specifically for temperature or dew point. The combination of forecast errors of both variables may offset when converted within the icing algorithms to relative humidity.

As a final note, the errors associated with wind speeds, both u and v, are generally small, especially when we remember the increasing magnitude of the observed wind speeds at the higher levels.

While the bias and RMSE calculations show some indication of the quality of ARPS forecasts of the upper atmosphere, the bias numbers do not give us an "acceptance" measure for using these forecasts in the generation of products and AIVs for the Tinker BWS. In a study to evaluate the usefulness of various mesoscale model forecasting systems to United States Air Force theater numerical weather prediction needs (Cox et al.1998), threshold criteria were identified, shown in Table 4.2, to provide an upper and lower limit on what forecast errors are considered acceptable using mesoscale model output. Originally, dew point depression criteria were specified versus actual dew point values. For consistency reasons with the surface bias and RMSE calculations it was decided to do a direct dew point accuracy evaluation. As an additional measure of the agreement of the ARPS forecasts to the specified error range listed, the number of forecasts (each grid point within the specified verification level) that met these categories were determined as a percent of the total. The number of forecasts within the desired accuracy range that had positive biases or negative biases were also determined separately. Table 4.3 and Figure 4.41 give the results of this process for the same evaluation variables as in Table 4.1.

Temperature in the lower levels only meets the accuracy criteria up to 50% of the time for the surface and 850 mb levels, and drops significantly by the end of the forecast period. The amount increases to above 80% for the 700mb and higher levels, with no significant drop for the later forecast times. The dew point temperatures are "accurate"

Parameter	Accuracy Threshold
Temperature	± 2 °C
Dew point	± 2 °C
Wind Speed:	
less than 10 m s^{-1}	± 1 m s^{-1}
greater than 10 m s^{-1}	± 2.5 m s^{-1}
Wind Direction	± 30 °
Sea level pressure	± 1.7 mb

Table 4.2: Desired forecast accuracy (as noted in Cox et al. 1998, except for dew point)

			02hr (15UTC)			05hr (18UTC)			08hr (21UTC)		
Level		Units	%	% +	% -	%	% +	% -	%	% +	% -
Surface			46.4	61.7	42.5	36.3	51.6	53.8	30.4	50.9	55.5
850mb			54.3	27.5	76.0	56.3	30.2	73.2	50.1	30.1	73.7
700mb	Temperature	° C	84.3	21.7	80.6	86.8	28.1	74.2	83.9	27.7	74.6
500mb			97.8	35.4	66.6	95.1	43.9	58.1	91.9	52.4	49.8
300mb			91.7	38.6	63.5	91.4	49.7	52.4	87.3	53.6	48.6
Surface			61.3	71.8	31.4	43.2	74.7	29.7	35.4	74.5	31.0
850mb			47.5	60.6	43.5	49.5	63.3	40.6	48.5	58.5	45.5
700mb	Dewpoint	° C	38.8	61.2	43.8	42.7	62.6	42.0	44.4	61.1	43.2
500mb			64.3	54.0	50.5	64.0	54.2	50.4	66.9	52.8	51.8
300mb			94.6	38.1	63.9	93.6	49.2	52.9	90.2	53.0	49.2
Surface			68.9	51.2	48.8	60.5	65.8	34.2	59.4	68.0	32.0
850mb	Wind		82.0	52.8	49.6	78.8	61.8	40.7	75.7	67.8	34.8
700mb	Direction	°	94.1	53.4	48.7	94.2	54.1	47.9	91.6	55.9	46.2
500mb			99.2	46.9	55.1	97.8	49.5	52.4	96.3	56.5	45.5
300mb			99.4	56.3	45.6	98.2	62.0	40.0	96.3	64.4	37.6
Surface			48.4	45.5	54.5	42.0	45.4	54.6	37.1	41.8	58.2
850mb	Wind		52.2	46.4	57.3	46.7	53.4	50.7	41.6	54.0	50.7
700mb	Speed	m s^{-1}	57.3	46.1	57.3	56.8	44.5	58.9	54.4	43.7	59.9
500mb			67.0	50.0	52.9	62.0	52.7	50.4	54.4	48.5	55.1
300mb			52.0	51.6	52.1	45.0	48.9	55.4	38.3	50.2	54.8
Surface	SLP	mb	55.5	61.9	41.6	61.6	51.6	51.6	43.3	68.7	35.8

Table 4.3: WOP97 average field percentage within accuracy criteria for surface, 850mb, 700mb, 500mb, and 300mb pressure levels for each desired quantity at the 02, 05, and 08 hour forecast points (15, 18, and 21 UTC respectively). Includes total percent within accuracy criteria (%), % total within accuracy criteria with a positive bias (%+), and the % total within accuracy criteria with a negative bias (%-).

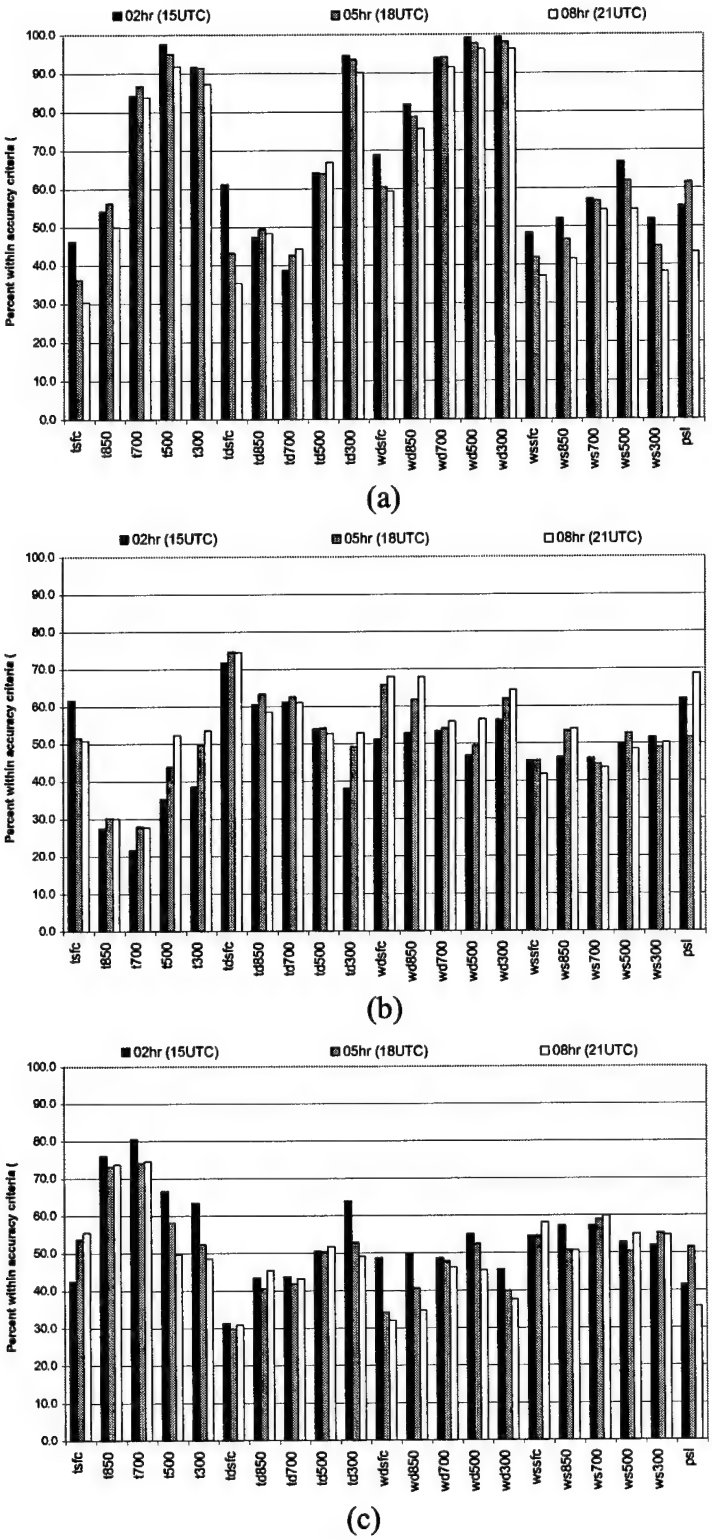
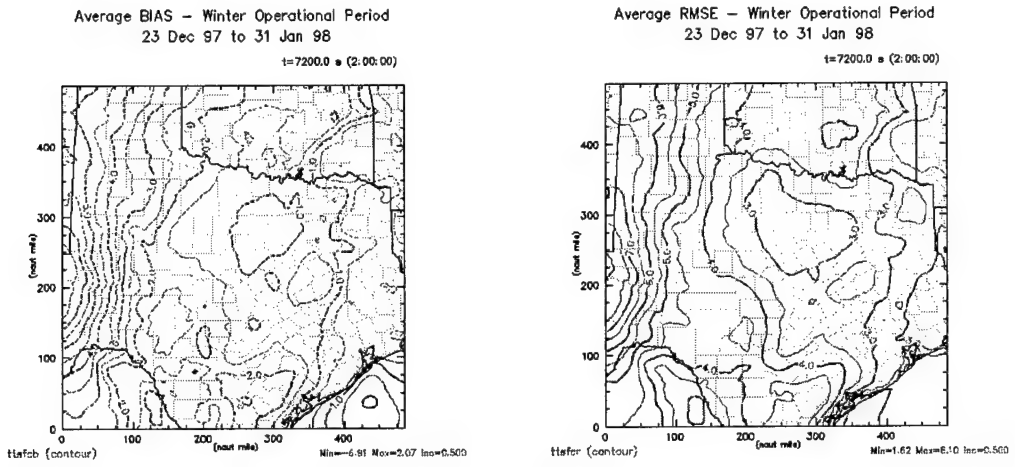
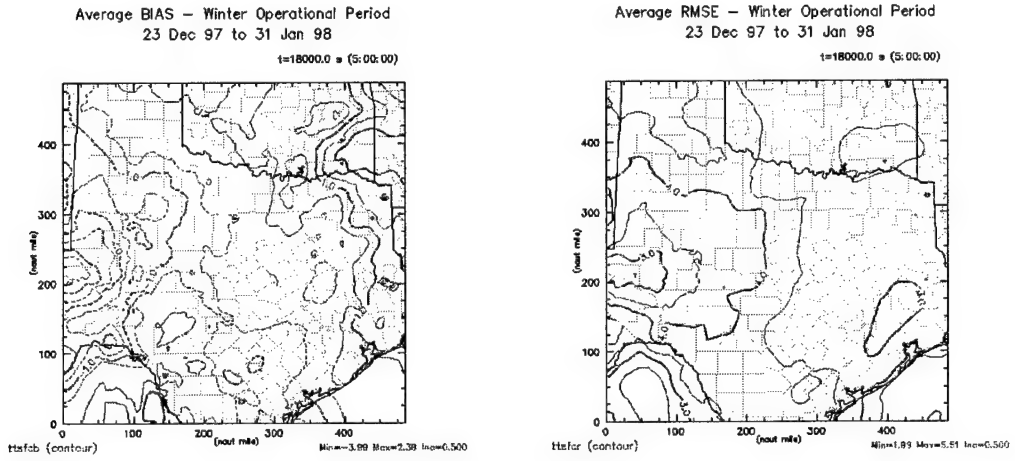


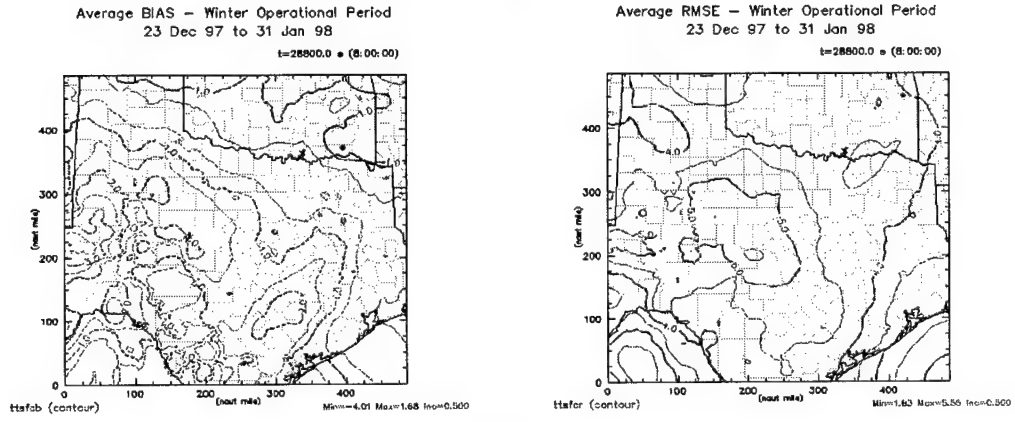
Figure 4.41: Data in Table 4.3 in graphical format. Includes (a) total percent within accuracy criteria (%), (b) % total within accuracy criteria with a positive bias (%+), and (c) the % total within accuracy criteria with a negative bias (%-).



(a)

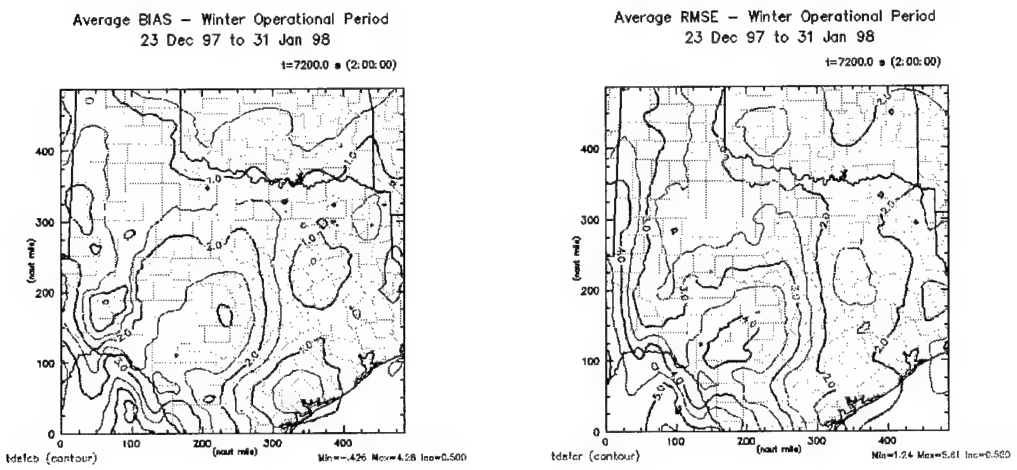


(b)

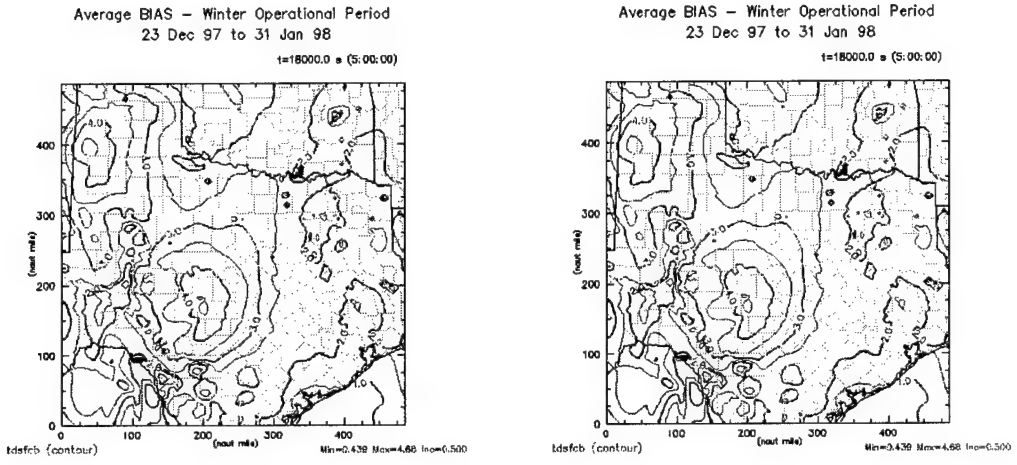


(c)

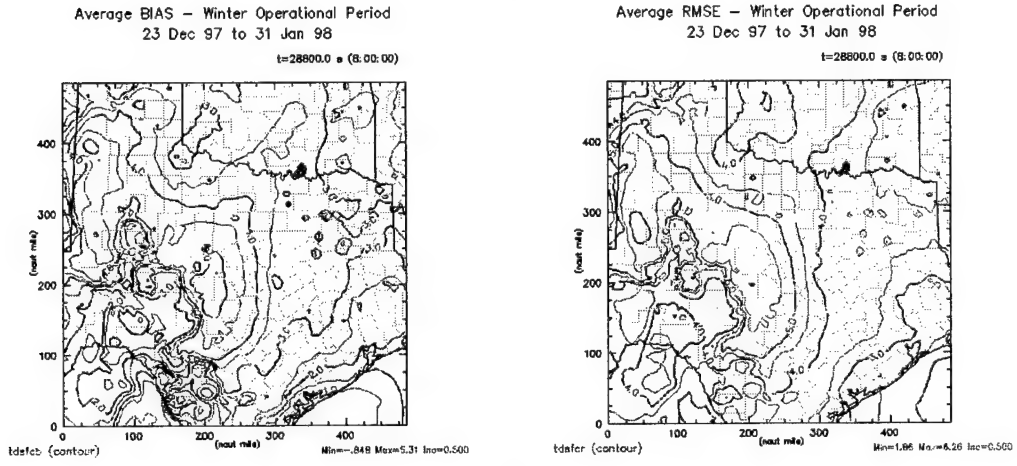
Figure 4.42: Average bias (left) and RMSE (right) for surface temperature (°C) for (a) 02, (b) 05, and (c) 08 hour forecast point (15, 18, 21 UTC respectively). (Negative values are indicated by dashed isopleths, while solid ones indicate positive values)



(a)

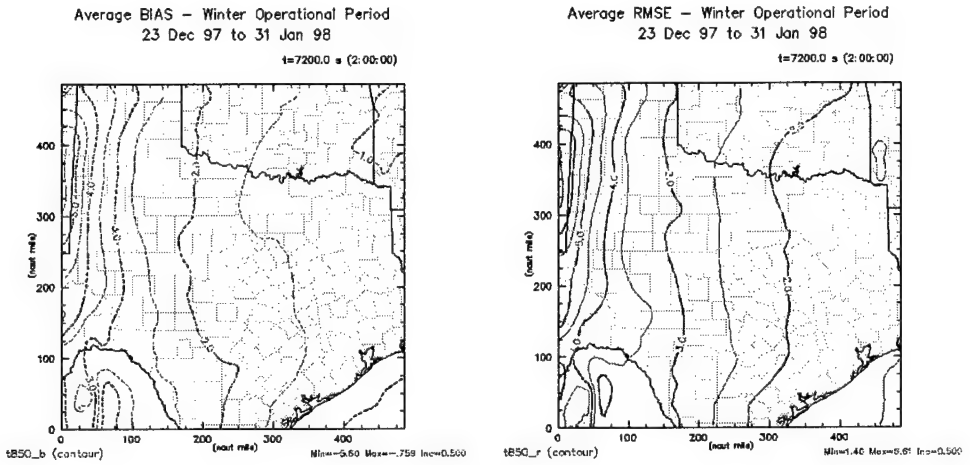


(b)

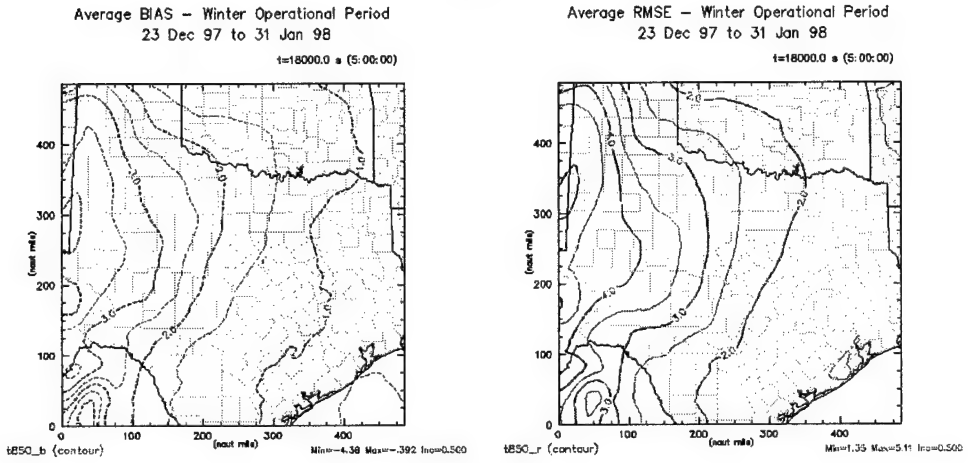


(c)

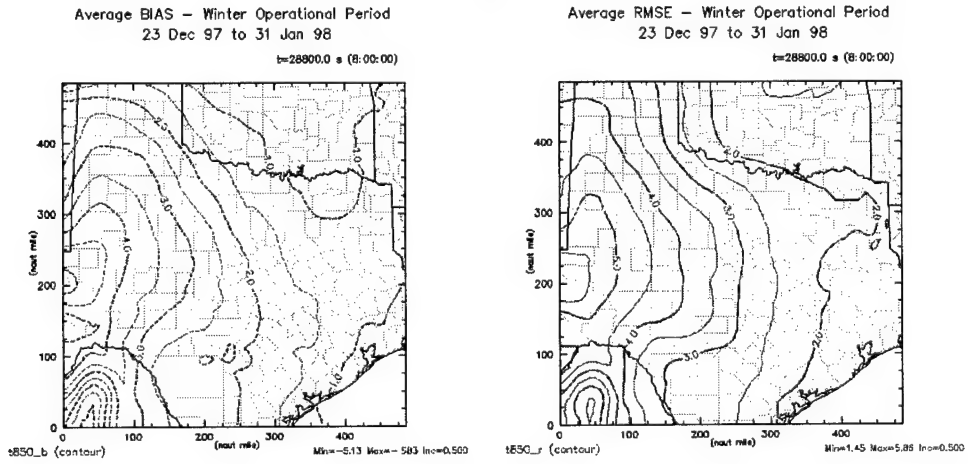
Figure 4.43: Same as Figure 4.42 except for surface dewpoint (°C).



(a)

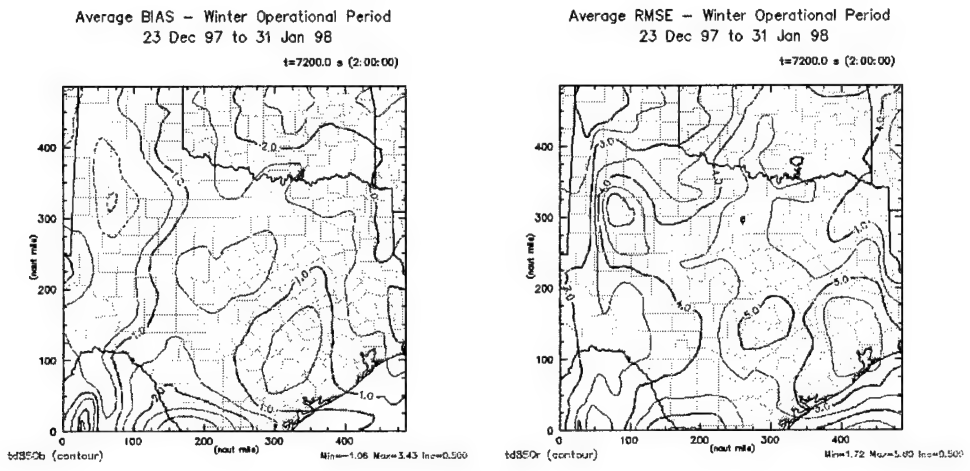


(b)

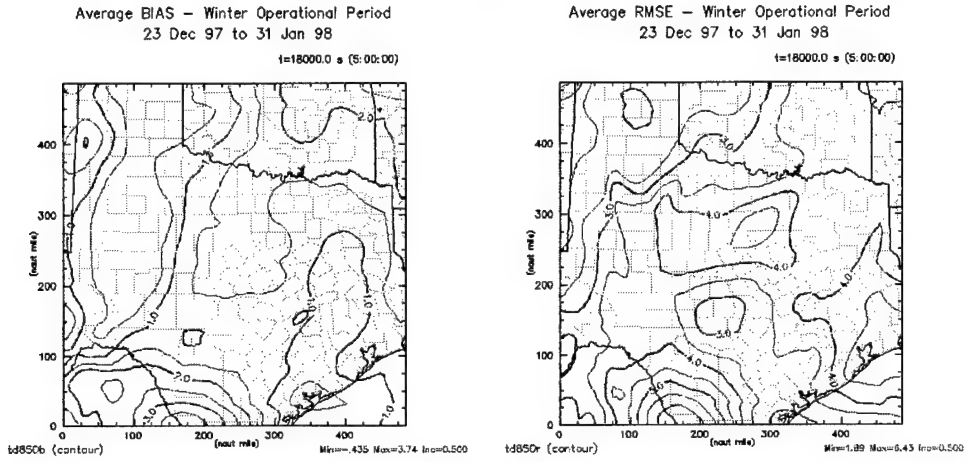


(c)

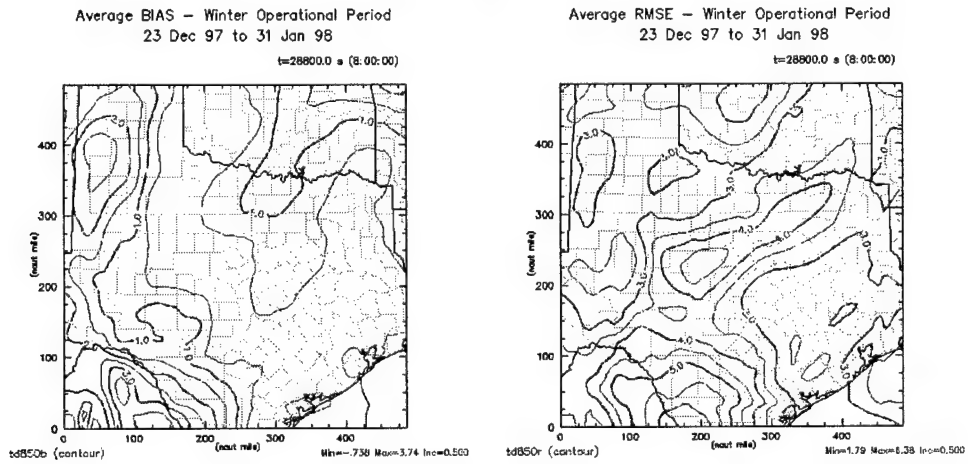
Figure 4.44: Same as Figure 4.42 except for 850mb temperature (°C).



(a)



(b)



(c)

Figure 4.45: Same as Figure 4.42 except for 850mb dew point (°C).

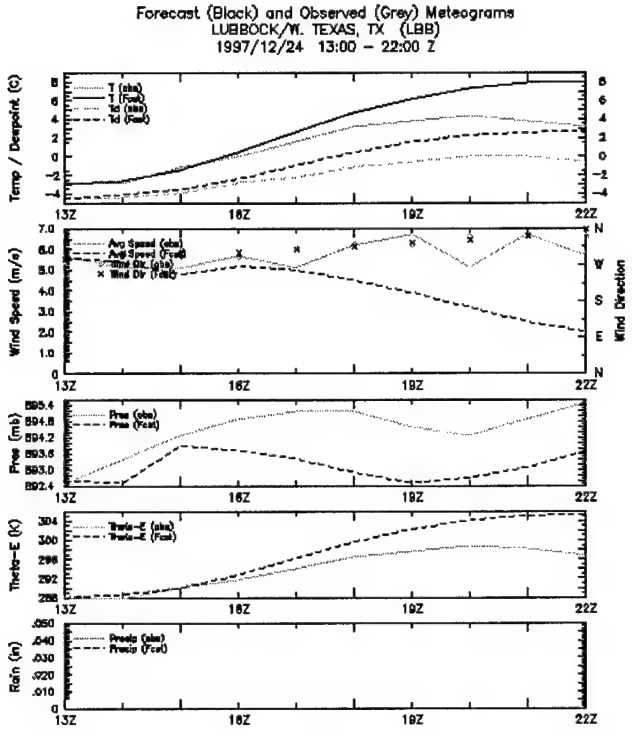
even less often higher in the domain, with only 40% of the 02 hour (15 UTC) data points in the desired range at 500mb. The wind speed forecasts only meet the listed accuracy criteria about half the time, which remains consistent for all levels and forecast hours.

To get a feel for the spatial characteristics of the forecast error measurements, the grid point errors were averaged, for each forecast hour, at each of the five pressure levels. The only significant spatial feature noted in examinations of these plots was a general disparity between the western and eastern halves of the forecast domain in the low level temperature and dew point fields. Figures 4.42 and 4.43 show the average grid point bias and RMSE results for the surface and 850mb level temperature and dew point, forecasts for the entire horizontal domain. We see that the average biases indicated in Table 4.1 are not uniform across the domain. The 02 hour surface temperature forecast errors are around –1 °C in the eastern half of the domain and increase to above –5 °C at the western edge. By the 08 hour forecast point, the errors become slightly more uniform, but still increase in magnitude from east to west. The same relationship is shown in the dew point plots except that the transition region is much smaller and located more toward the central portions of the domain. The trend is still evident in the 850mb temperature and dew point plots, Figures 4.44 and 4.45, but are smaller in the disparity of the magnitude of the errors between the domain halves. The 850mb trend is easier to rationalize since the western portion will be in closer proximity to the ground than the eastern half due to the sloping terrain. This, however, does not explain the presence of the same features for the surface based calculations. The surface errors may be indicative of the soil moisture and vegetation difficulties ARPS has along with the inherent errors when evaluating boundary layer processes, as well as initialization errors. We can return to the surface

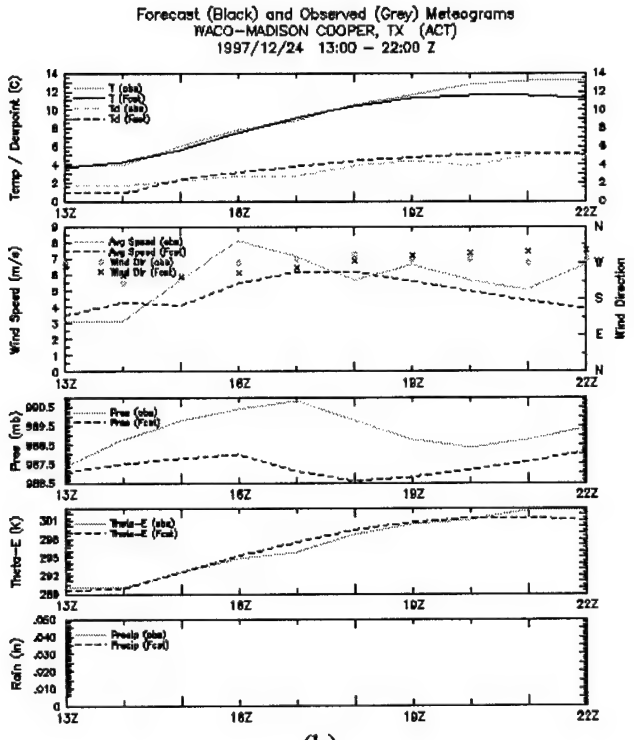
observation verification discussed earlier to see examples that illustrate these same spatial differences in the surface forecast errors. Figure 4.46 shows the actual forecast and observed meteogram comparison for Waco and Lubbock, TX, on 24 January 1998. Here we see that the forecast at Lubbock (west half of the domain) is too warm and moist while the forecast at Waco (east half) is correctly indicating the moisture trend but is too cold in its temperature forecast.

The spatial characteristics of the forecast error at the upper levels, 500 and 300mb (not shown), showed no significant variation across the domain. The magnitudes and trends indicated by the overall averages in Table 4.1 were consistent across the domain.

In summary we see that the predominant errors in the ARPS forecasts are within the lower layers of the model domain, possibly due to various boundary layer factors which the ARPS model does not treat adequately. The vertical extent of the dew point errors is much greater than that for temperature. The surface observations and the surface data extracted from the RUC analysis fields are in relative agreement, which may indicate that the presence of this moisture bias extending further up into the atmosphere than expected is reasonable. We also see that the greatest errors appear to be associated with regions near the boundaries of the model domain. This agrees with the features identified by Warner (1997) in that the quality of mesoscale model forecasts may be significantly hindered by performance at the lateral boundaries. In general, however, the forecasts in this work are not in such disagreement with the observations or significantly outside the accuracy criteria to make a verification of the icing forecasts pointless.



(a)



(b)

Figure 4.46: Meteogram comparison for forecast on 24 December 1997 valid at (a) Lubbock, TX, and (b) Waco, TX.

Section 4.3 Verification of Icing Forecasts

We now can proceed into the evaluation of the icing forecasts generated by COMET-Tinker with some degree of confidence that the forecast data used as input into the algorithms are fairly accurate. In the context of the methodology outlined in Chapter 3, we must remember that the measurement of the ability of each of the seven algorithms to accurately predict the presence of icing is not a simple procedure. Instead of determining whether the forecast was correct at each point within the model, we can only examine the locations where a report of the icing conditions was explicitly made. Secondly, the very nature of those reports does not allow us to use the common accuracy measures, such as the false alarm rate and skill scores, associated with categorical forecast verification. In this manner, we are not measuring each forecast algorithm's quality in a complete sense, but rather their comparative abilities to agree with conditions reported at a set of specific points in time and space. To get a thorough understanding of the relationships between the algorithms, we wish to examine not only their average ability to predict instances of observed icing over the entire set of forecasts, but to see if trends and patterns exist when a separation based on the various valid hours and individual forecast days is done. Secondly, we realize that to the aviation community the presence of the more severe icing intensities, moderate or greater, is more critical to flight operations and safety than those in the lower intensities, trace to light. An algorithm which can effectively capture the greater portion of the more severe events is in this respect the most desirable.

The average detection rates of "Yes" and "No" reports, POD_{yes} and POD_{no}, using all icing reports received are the basic measure of the abilities of the icing algorithms. These results averaged over the entire WOP97 forecasts are given in Figure 4.47. All

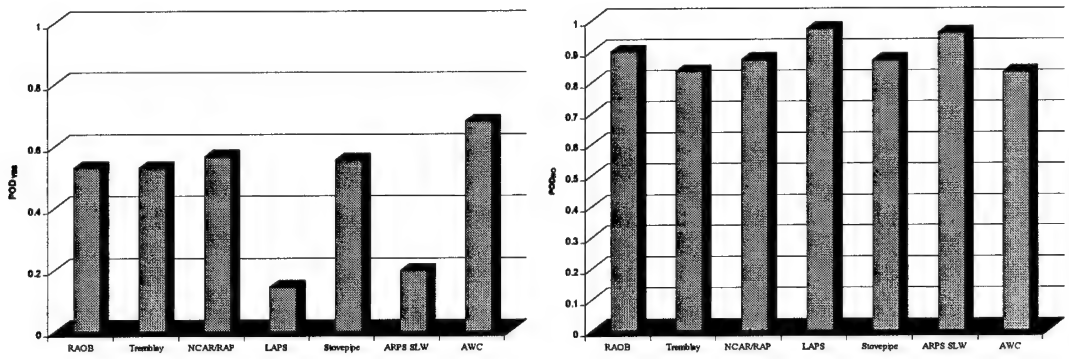


Figure 4.47: Mean (a) POD_{yes} and (b) POD_{no} for all PIREPs over the entire WOP97.

algorithms were able to accurately predict the majority of "No" icing events. This is because of the relative ease in which any forecasting method can pick up strong indications of the absence of icing. For days in the WOP97 where the majority of reports were negative, forcing mechanisms for icing were so completely absent that the majority of algorithms had little or no icing conditions forecast. While not trivial, the ability to forecast "No" icing is generally considered easier than making a "Yes" determination. Single input indicators to the algorithms, such as temperatures well above freezing, provide little doubt as to what the forecast should be. In attempting to actually predict icing conditions, especially regions of trace or light icing, the complex combination of algorithm inputs make it harder for the algorithms to accurately forecast the "Yes" events, especially given the fact that the model output of variables used by the algorithms is not perfect. Thus the POD_{yes} values are expected to be lower. The best detection rate of "Yes" icing reports was attained by the AWC algorithm at just above 60%, with the

majority of algorithms only able to detect less than half of the icing events sampled. We see that the LAPS and ARPS SLW algorithms caught a significantly smaller portion of the icing occurrences than the remaining five. The similarity of the physical rationale of the RAOB, NCAR/RAP, Stovepipe, and AWC algorithms can be used to explain their general agreement. It is interesting to note, however, that the Stovepipe algorithm, which is supposed to be an "improvement" of earlier versions of the NCAR/RAP methodology, did not generate a better overall detection rate. On initial examination, the Tremblay algorithm, which is based more on microphysical parameterizations used in numerical weather prediction than on synoptically-based physical reasoning, appears to do just as well as the others. Figure 4.48, however, which gives the mean impacted area (IA) and volume (IV) measurements of the various algorithms indicates that the relationships noted in the detection rates may not represent the relative effectiveness of the algorithm due to the significantly different areas (or volume) where icing was

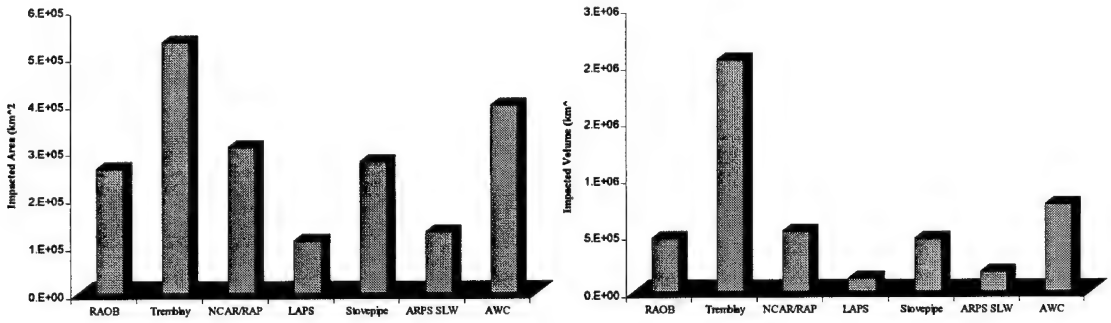


Figure 4.48: Mean (a) Impacted Area (km^2) and (b) Impacted Volume (km^3) for entire forecast period.

forecast. On average, the impacted area (volume) of the Tremblay algorithm covers nearly 60% (15%) of the horizontal (three dimensional) domain. This is significantly more than the other algorithms which have impacted areas covering only 10-50% of the

domain, and impacted volumes encompassing only 5-10% of the total volume. The fact that the LAPS and ARPS SLW algorithms forecast the smallest spatial extent of icing within the forecast domain may explain the relatively low detection rates of "Yes" events they exhibit. The small spatial extents of these forecasts is likely due to the insufficient ability of the ARPS model to correctly determine the liquid water content profile of the upper atmosphere. In regards to the effectiveness of these two algorithms, they may only be able to assess the most "strongly forced" icing events which other routine forecasting techniques and analysis methods available to the duty forecaster will also indicate.

The resulting area and volume efficiencies calculations (which represent the detection rate per unit area and volume respectively) are shown in Figure 4.49. In general the best efficiencies are exhibited by the "threshold" algorithms, with the exception of LAPS. In other words they do the best at maximizing their detection rates while keeping the extent of the forecast icing areas to a minimum. One possible explanation

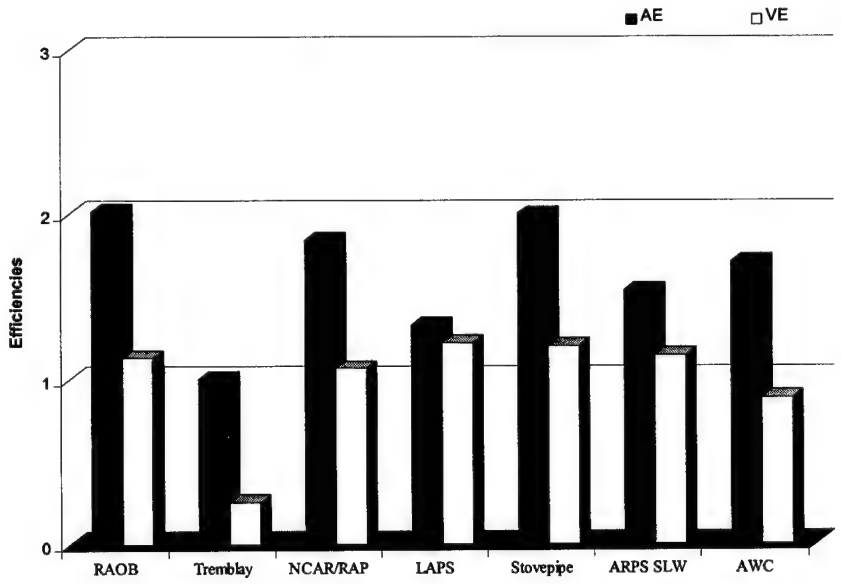
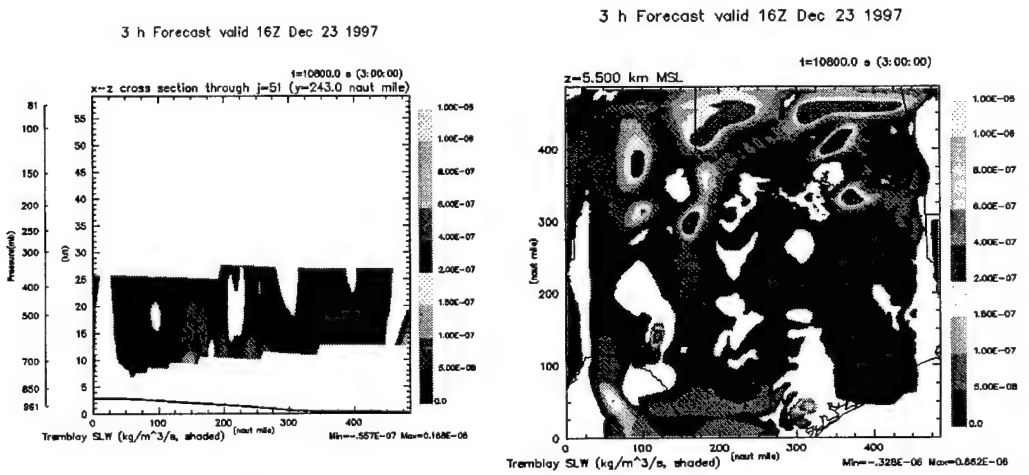


Figure 4.49: Mean Area Efficiency (AE; $x10^{-6}$ km^{-2}) and Volume Efficiency (VE; $x10^{-6}$ km^{-3}) for entire forecast period.

for the LAPS algorithm lagging the others is the fact it must rely on the ARPS water variables, q_c and q_i, to assess the presence of clouds. The general agreement between LAPS and ARPS SLW detection rates and impacted regions supports this. The relatively low POD_{yes} associated with the LAPS (and also the ARPS SLW) algorithm detracts from the efficiency similarities with the RAOB, AWC, and NCAR.RAP algorithms, in that they may only be efficient by not forecasting icing unless "strongly forced" as previously mentioned. In the case of the Tremblay algorithm, we see that the extremely high impacted area and volume amounts causes its efficiency to be much lower that the other algorithms, even with its high detection rate. The volume efficiency exhibited by the Tremblay algorithm is especially poor. The reason for this can best be understood by examining a typical icing forecast made by the algorithm. Figure 4.50 shows a pair of Tremblay icing forecast plots from 23 December 1997. The first panel (a) is a north-south cross-section showing the vertical extent of the forecast icing region. The second (b) gives a horizontal slice roughly through the middle of the icing region



(a) (b)

Figure 4.50: Tremblay icing forecast at (a) a north-south cross-sections through the center of the domain and (b) an altitude of 18,000 feet (FL180) on 23 December 1997. (Any shading (value > 0) indicates a "Yes" icing forecast)

indicated in (a) and is generally representative of all horizontal slices throughout the vertical on that day. From these two images we see that the algorithm is essentially forecasting icing in the entire horizontal domain between the temperature range it associates with icing, 0 to –40°C. It is expected that even a perfect set of forecasts in this case would still show low efficiencies due to the extent of the icing region forecast.

As with the surface and upper air verification, isolation of the hourly and daily forecasting characteristics of the algorithms can be useful to see situations where one algorithm performs better than the others. We desire to see a consistent forecast ability across all daily runs, as well as no significant drop in quality as the forecast lead time increases.

Figures 4.51 through 4.54 show the average POD_{yes}, IA, AE, and VE for all icing forecast algorithms for every forecast hour. The POD_{yes} at each forecast hour is consistent with the averaged results above. The LAPS and ARPS SLW consistently make more incorrect forecasts than the other algorithms. There is a general decreasing trend in POD_{yes} over the temporal extent of the forecasts. There is, however, a noticeable drop at the 04 hour forecast point for all algorithms. It is not completely understood why this occurs. Figure 4.52 indicates that on the average there is little change in the spatial extent of icing areas forecast throughout the nine hour period. This lack of variability agreed well with an examination of the actual sets of daily forecast plots that were displayed in real time to duty forecasters. An examination of the spatial coverage and intensity distribution of the PIREPs at the various valid times also showed no significant fluctuations. Thus the reports used each hour were not sampling drastically different areas of the domain, nor were there more trace reports, theoretically making things harder

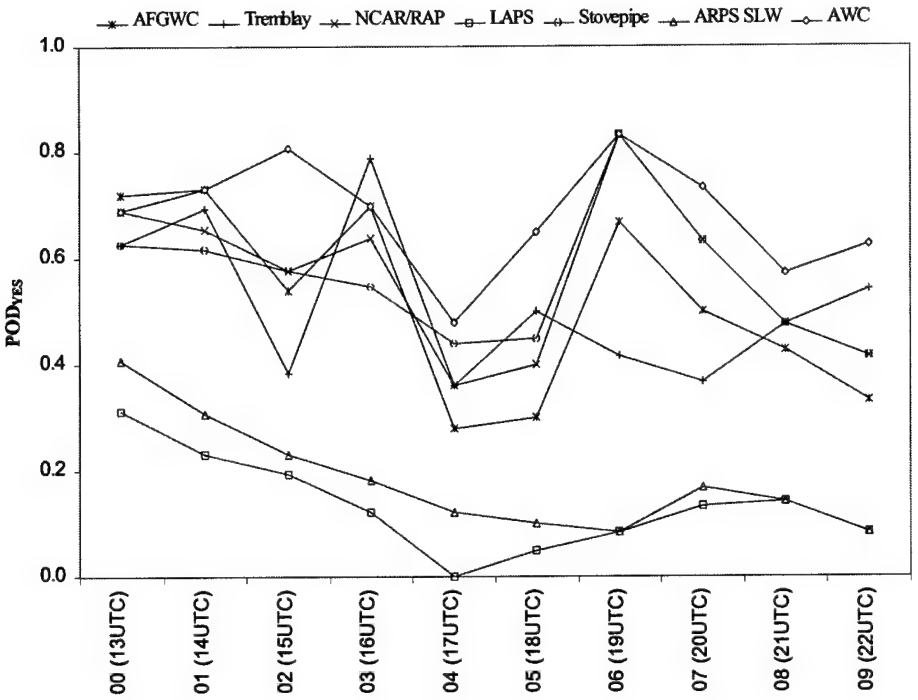


Figure 4.51: Average POD_{yes} for all icing forecast algorithms at each forecast hour.

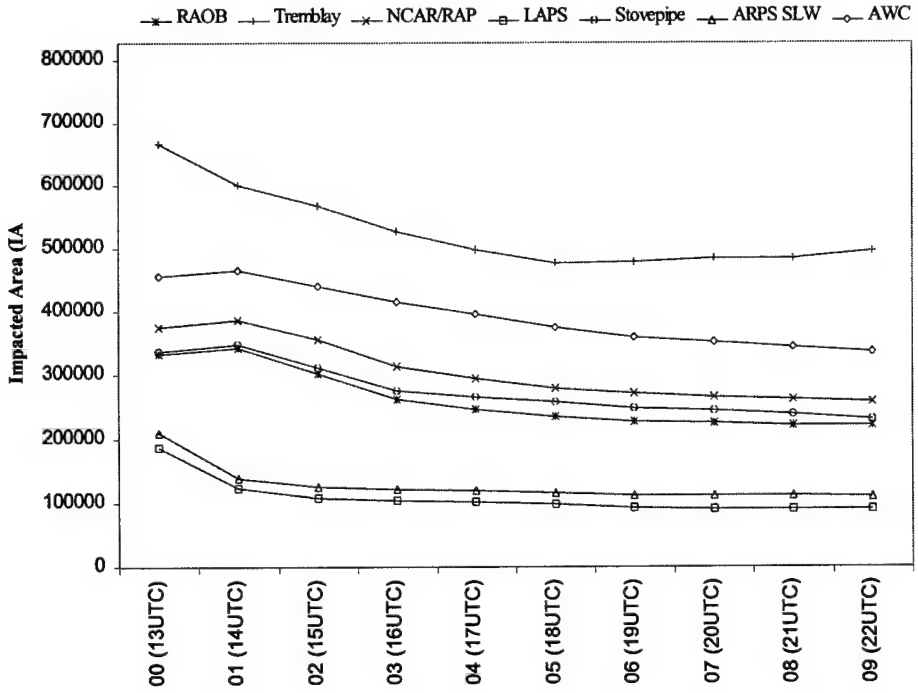


Figure 4.52: Average Impacted Area (IA) for all icing forecast algorithms at each forecast hour.

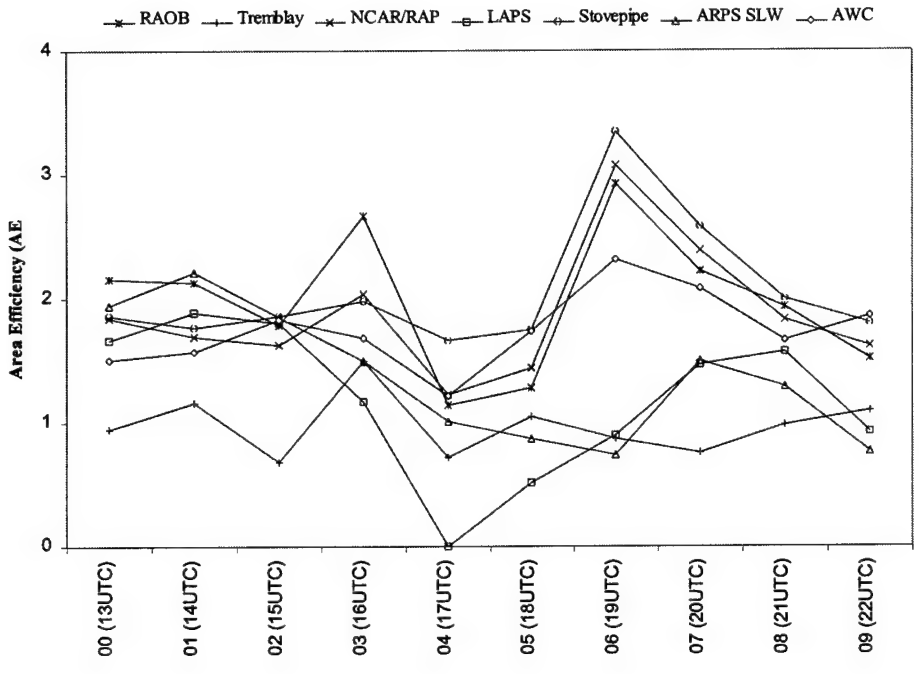


Figure 4.53: Average Area Efficiency (AE) for al icing forecast algorithms at each forecast hour.

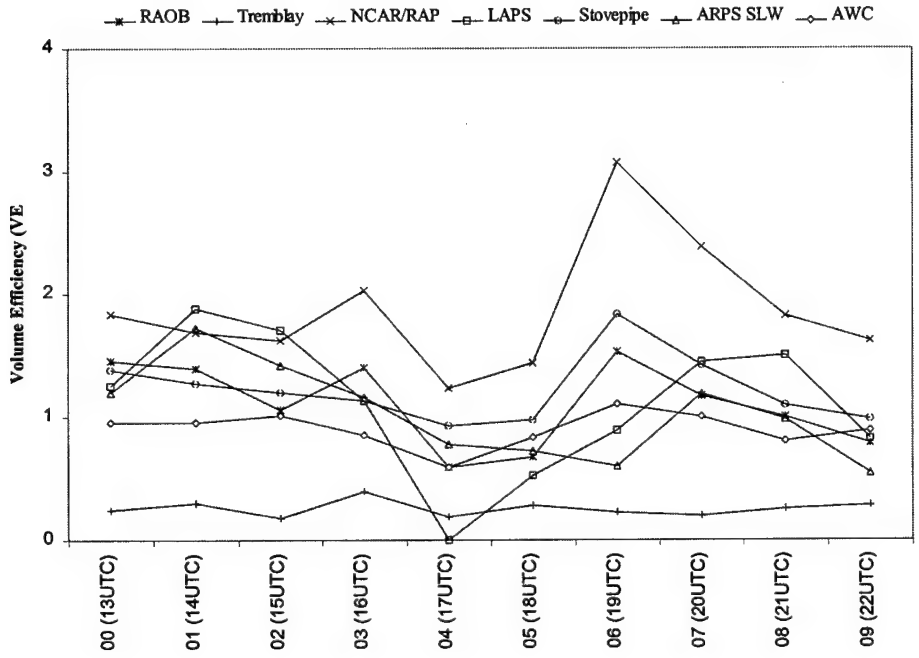


Figure 4.54: Average Volume Efficiency (VE) for al icing forecast algorithms at each forecast hour.

on the algorithms to pick up the icing events. However, Figure 3.5.(b) does show a decrease in the number of observed "No" reports after 16 UTC. The elimination of these, easier to forecast, "No" reports may result in the drop in POD_{yes}. However, there is no increase in "No" reports, corresponding to the increase in detection rate at the 06 hour (19 UTC) forecast. The change may also be due to the limited number of PIREPs available each hour, averaging 2-4 reports, which may just mean that natural variability in the detection rates obtainable is being observed. Verification over a longer operational period may eliminate this feature by providing a larger set of PIREPs to verify.

The efficiency scores do not indicate that any particular algorithm is routinely better than the others over all forecast times. Again, the small spatial extent of the LAPS and ARPS SLW forecasts make them appear equally as efficient as the others, while the excessive nature of the Tremblay forecasts drops it to the bottom of the list.

It was evident from Figure 3.5 that some daily variability in the number and intensity distribution of the reports received was present. Table 4.4 lists the daily POD_{yes} values for each algorithm on days where at least one "Yes" report of icing was received. The detection rates exceeding 50% are shaded gray, while those days with values below 20% are indicated in the ***bold*** lettering. The LAPS and ARPS SLW have predominantly similar, and relatively low, daily results. The AWC forecasts are consistently the best, with more than half the events captured on any given day. In Figure 4.47 we saw the other threshold algorithms had only slightly lower average detection rates. There is, however, a greater number of days where each was not able to get the detection rate to the 50% mark. In fact, the relatively high POD_{yes} for the RAOB, NCAR/RAP, and Stovepipe forecasts on 24 and 26 December 1997, days with considerable numbers of

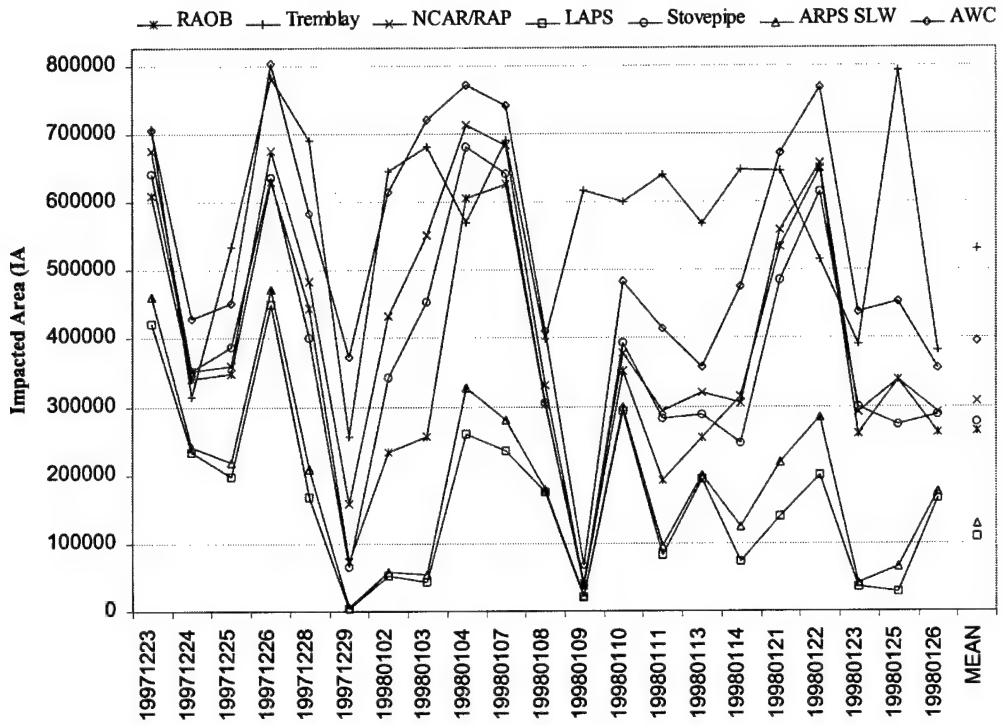
	RAOB	Tremblay	NCAR/ RAP	LAPS	Stovepipe	ARPS SLW	AWC
19971223	0.61	0.79	0.67	0.27	0.61	0.42	0.70
19971224	1.00	***0.18***	0.91	0.36	0.91	0.36	1.00
19971225	***0.20***	***0.20***	***0.20***	***0.00***	***0.20***	***0.00***	0.40
19971226	1.00	0.84	0.95	0.37	0.84	0.37	1.00
19971228	0.45	0.76	0.48	***0.10***	0.48	***0.14***	0.55
19971229	***0.00***	***0.00***	***0.00***	***0.00***	*0.00*	***0.00***	1.00
19980102	0.33	0.33	0.67	***0.00***	1.00	***0.00***	1.00
19980103	0.50	0.50	***0.00***	***0.00***	***0.00***	***0.00***	0.50
19980104	0.42	0.17	0.67	***0.17***	0.67	***0.17***	0.75
19980107	0.56	0.56	0.60	***0.07***	0.58	***0.19***	0.63
19980108	0.56	0.28	0.44	***0.06***	0.44	***0.06***	0.67
19980109	***0.00***	***0.00***	***0.00***	***0.00***	***0.00***	***0.00***	***0.00***
19980110	0.78	0.67	0.78	0.56	0.78	0.56	0.89
19980111	***0.00***	1.00	1.00	***0.00***	1.00	***0.00***	1.00
19980113	***0.00***	***0.00***	***0.10***	***0.00***	***0.20***	***0.00***	***0.20***
19980114	***0.00***	0.33	0.33	***0.00***	0.33	***0.00***	0.67
19980121	0.25	0.25	0.75	***0.00***	0.50	***0.00***	0.75
19980122	0.35	0.35	0.35	***0.00***	0.53	***0.12***	0.65
19980123	0.23	0.54	***0.15***	***0.00***	***0.15***	***0.00***	0.46
19980125	0.67	0.78	0.67	***0.00***	0.44	***0.00***	0.78
19980126	0.75	0.50	1.00	0.25	1.00	0.25	1.00

Table 4.4: Daily POD_{yes} for each forecast icing algorithm indicated. (Gray indicates POD > 50%, while ***bold*** is < 20%)

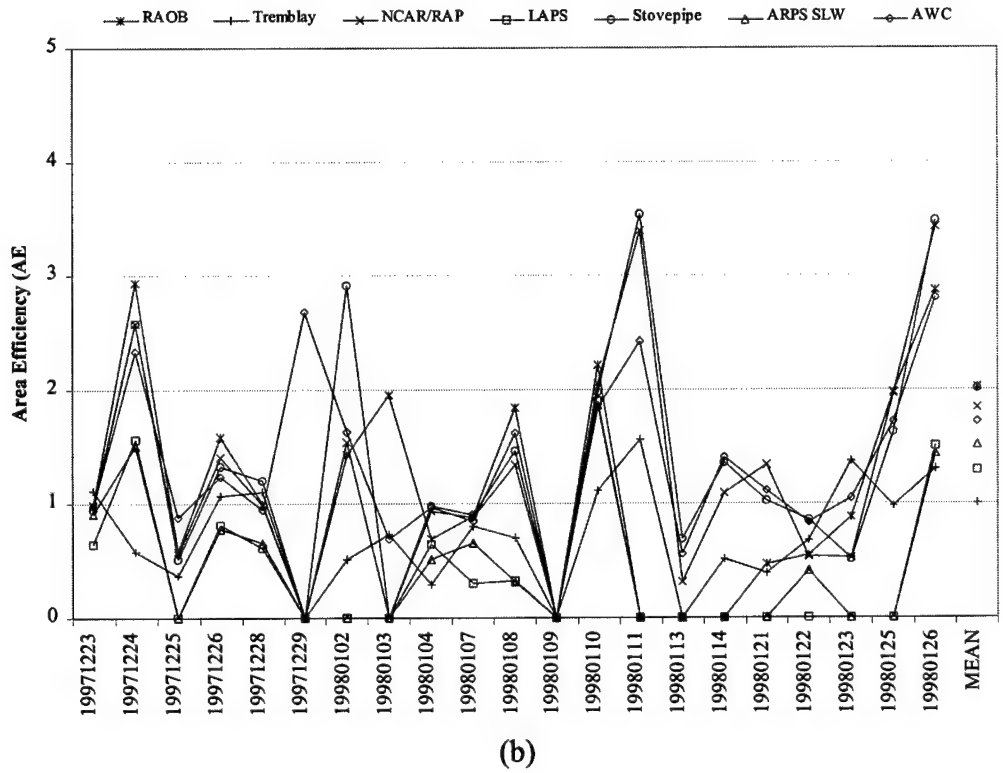
"Yes" reports (11 and 19 respectively) seems to compensate for days, with a considerable number of reports, where the detection rates are lower (i.e. 13, 22, and 23 January). All algorithms performed well on 10 January when icing forecasts of a "strongly" forced nature, associated with a relatively cold airmass moving through the northern part of the forecast domain, matched well with the observed reports (all in the northern third of the domain). Lastly, a few of the days had almost no detection ability evident from any of the forecast algorithms. These primarily coincided with days where small numbers of low intensity reports were present, making detection extremely difficult to do accurately.

The daily impacted area and volume determinations as well as the efficiency results for the daily forecasts are displayed in Figures 4.55 and 4.56. The daily IA and IV values show no significant variations from the WOP97 averages shown in Figure 4.48. In a subjective sense, comparing the relative maximums and minimums with the number of "Yes" icing reports , given in Figure 3.6, shows that the algorithms did forecast greater regions of icing extent on days where a larger numbers of icing reports were made. However, the efficiency results show that the larger icing regions forecast on those days did not necessarily imply an improved detection rate on those days. This may be a result of the spatial characteristics of the icing reports and not directly related to the changes in amount of icing area forecast.

One feature noted in the daily impacted area and volume characteristics in 4.55(a) and (c) was the elevated Tremblay levels in the middle of January. From 9 to 13 January the predominant forecast trend among all the algorithms, except Tremblay, was a reduction in the spatial extent of the icing forecast within the domain. This corresponded well with the strong forcing nature of the cold airmass in the northern part of the domain

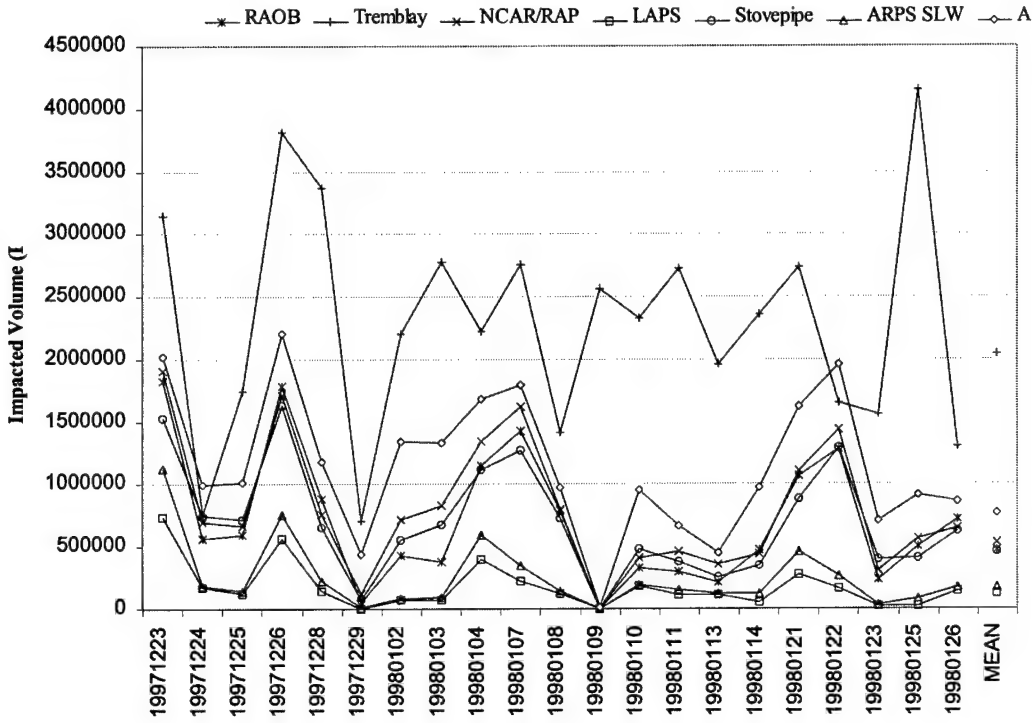


(a)

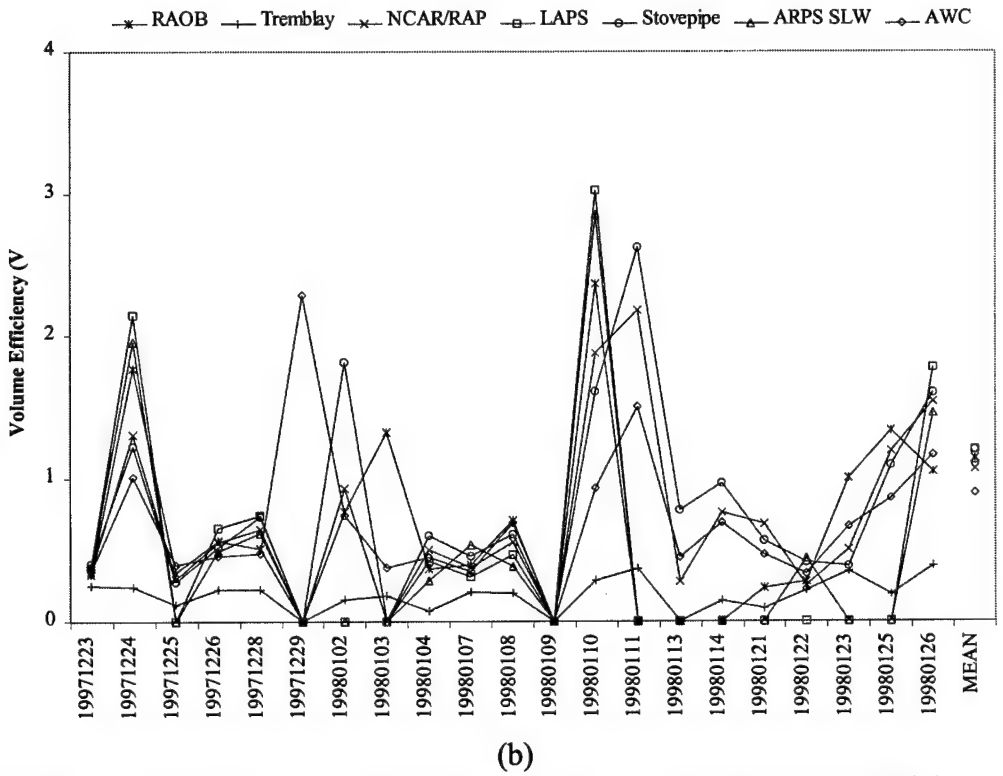


(b)

Figure 4.55: Daily (a) Impacted Area (IA) and (b) Area Efficiency (AE) encompassing all icing PIREPs



(a)



(b)

Figure 4.56: Daily (a) Impacted Volume (IV), and (b) Volume Efficiency (VE) encompassing all icing PIREPs

on the 10th, as well as relatively benign conditions experienced elsewhere in the region. The general consensus among algorithms gives the impression that the icing conditions experienced during this period were "easy" to forecast. Why then did the Tremblay forecasts include such a large portion of the domain? Figure 4.57 shows a comparison of the Tremblay and Stovepipe icing forecasts for 15 UTC (02 hour) on 10 January. The forecast of each at flight levels of 050 and 180 are displayed as well as the verification PIREPs valid at 14 and 15 UTC. The general area where icing was reported was in the northern portion of the domain, at lower altitudes. This matches well with both icing forecasts shown. The Tremblay forecast however includes additional icing regions in the central part of the forecast domain at higher altitudes which are not present in the other icing forecasts. Additionally, the limited sample of PIREPs indicated no real icing throughout the higher altitudes where the Tremblay method is forecasting it. The theory behind the formulation of the Tremblay algorithm says that, in the appropriate temperature range, anytime water vapor can change to cloud water faster than it can deposited on snow (wG-SDEP > 0) means that icing is possible. We see however that the use of the "anytime" threshold as has been presented, may be part of the reason behind the "overforecasting" and poor efficiency evident with the Tremblay algorithm. An increase in the threshold could yield an improved forecast result, and provide an assessment of when there is a "better" chance of experiencing icing.

However, the magnitudes of the SLW generation for the Tremblay algorithm are relatively narrow in range. While adjustments to the zero threshold could remove the erroneous icing regions in the upper levels, it could also reverse the correct forecast that was made in the low levels in the example shown in Figure 4.57. If changes were to be

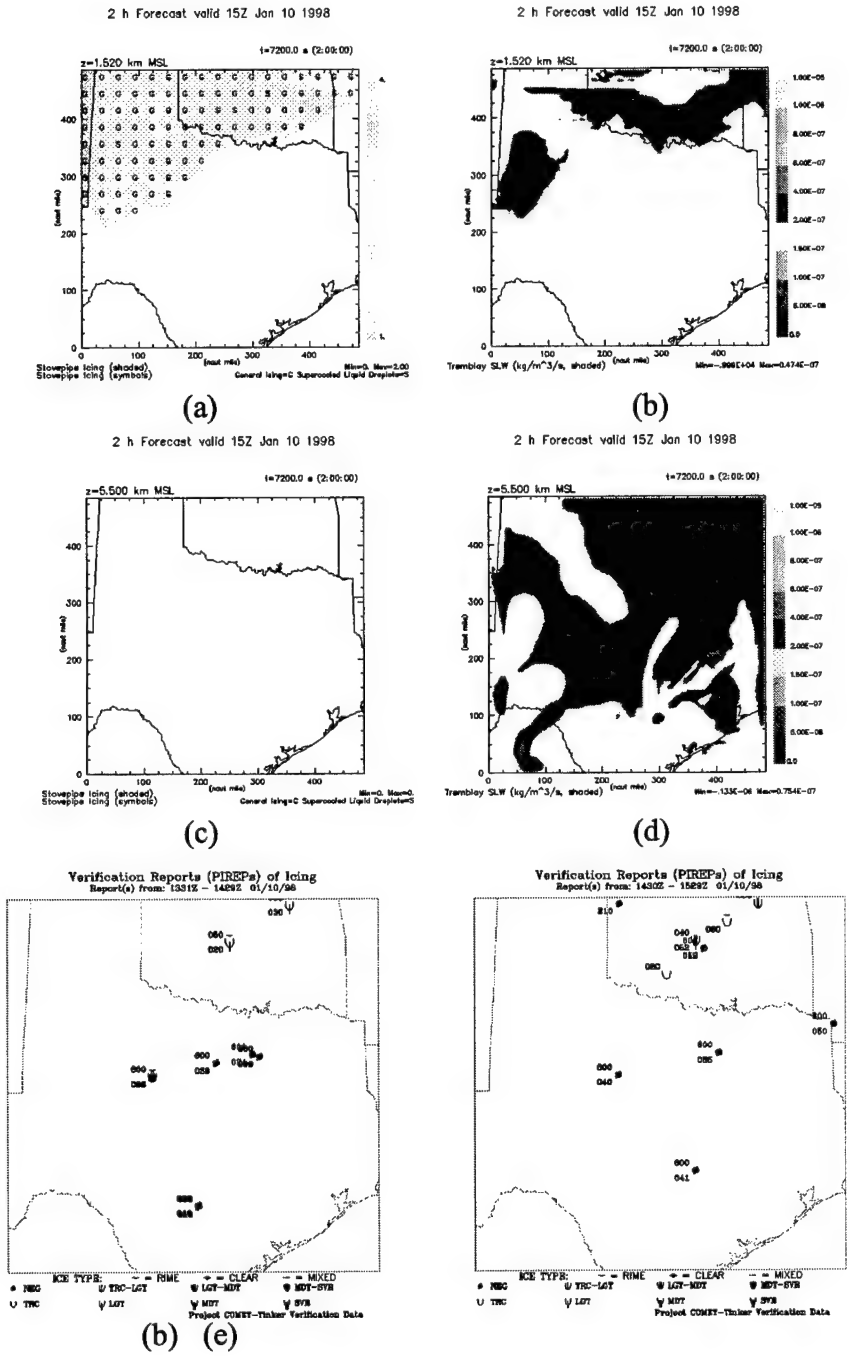


Figure 4.57: 02 hour (15 UTC) icing forecast on 10 January 1998, at altitudes of 15,000 feet (FL150) (a and b), and FL180 (c and d), for the Stovepipe (a and c) and Tremblay (b and d) icing algorithms. Corresponding (d) 14 UTC and (e) 15 UTC icing PIREPs for verification shown as well.

made in the thresholds of the Tremblay (or other algorithms) it would be hoped that the improvement would not increase the efficiency by decreasing the impacted region at the expense of the overall detection rate. In a second example, Figure 4.58 gives the 03 hour (16 UTC) RAOB and Tremblay forecasts at FL180 from 23 December 1997, along with the associated verifying PIREPs. The area in the northern part of the domain where both algorithms forecast icing corresponds well with the transmitted reports. This region in the Tremblay forecasts has larger SLW generation rates (SLW > $2x10^{-7}$ (kg m^{-3})s^{-1}) than the southeastern region where no reports of icing were indicated. Adjustment of the threshold could have reduced the extent of the forecast but we see that significant increasing of the value would push the forecast to the extreme low values of the LAPS and ARPS SLW results, where the forecast area would be so small that just a few correct forecasts/event pairs would give a good efficiency value. Figure 4.58 also points out a feature of the ARPS forecast that may explain some poor efficiency results of the ARPS SLW algorithm. The gridded nature of the forecast domain was evident in many of the ARPS SLW based icing forecasts in the form of the striations present in Figure 4.58(c). Here, the qualitative coverage of the forecast area generally agrees with the other forecast algorithms, while many individual points within the gridded forecast output do not. In other words, the sporadic nature of the icing forecast area can lead to a significant number of incorrect "missed" forecasts and may partly explain both the low POD_{yes} rate and the decreased IA and IV amounts displayed by the ARPS SLW algorithm. A second source of the striped regions may simply be from the reduction of the gridded forecasts to the desired constant height level. Figure 4.58(d) shows that this may not be true. Yet, in some regions where the forecast liquid water content is very close to zero, the

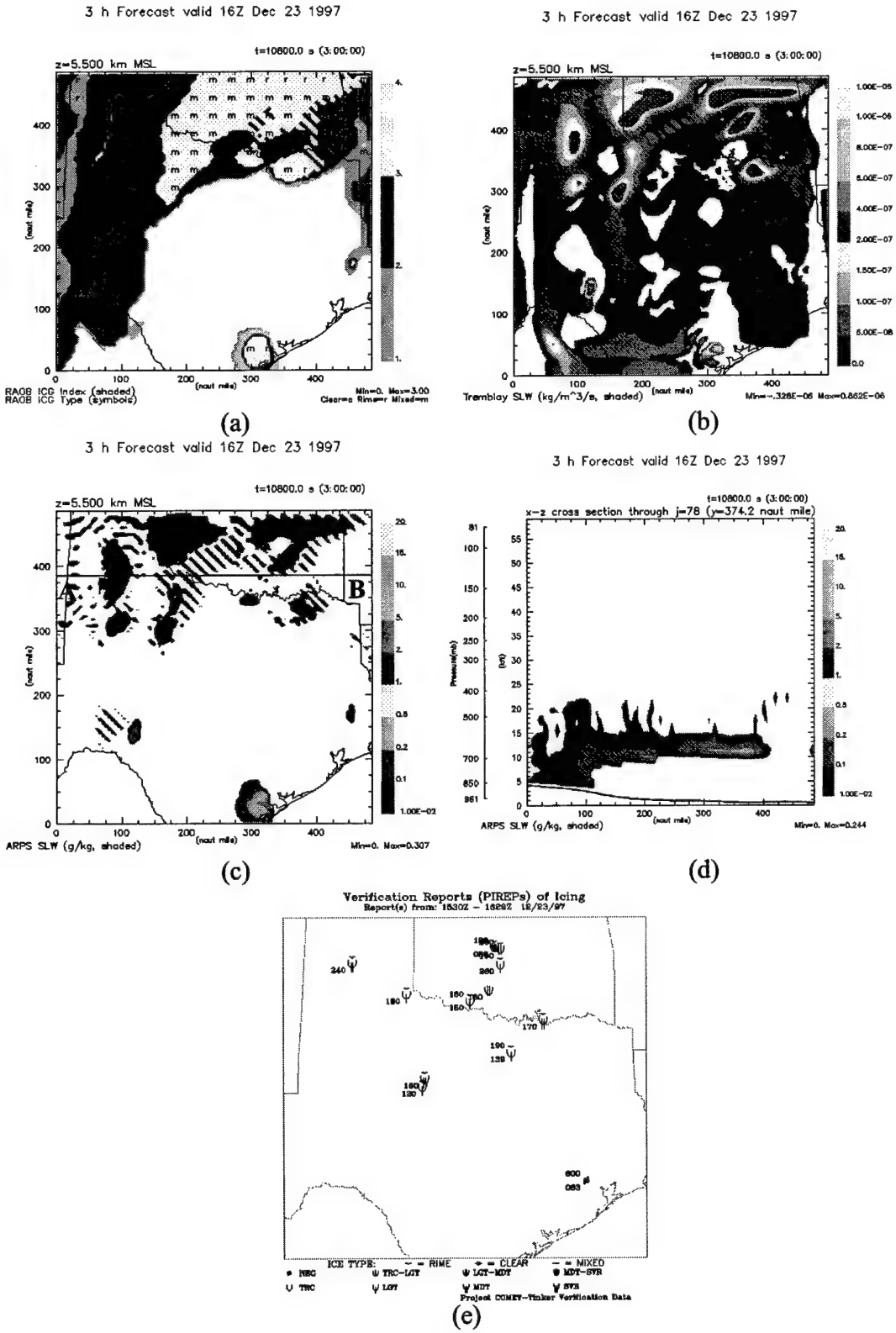


Figure 4.58: 03 hour (16 UTC) 23 December 1997, 18,000 feet (FL180) icing forecast for (a) RAOB, (b) Tremblay, and (c) ARPS SLW; (d) ARPS SLW east–west cross section (indicated by AB in (c)); and (e) the corresponding icing PIREPs.

interpolation to the desired level for the display may introduce the variations shown. This would not affect the verification against PIREPS since the nearest grid points are used and no interpolation is performed. If these stripes are a true feature of the forecast, it is believed that an ARPS SLW assessment that could incorporate those missing "stripes" into the forecast could improve the POD at a more significant rate than the increase in IA and IV. An overall improvement of the efficiency would result. Difficulties with the ARPS SLW algorithm are far more complex than just this. In other instances the limited amount of icing forecast by the ARPS SLW algorithm showed no such striped effect and was significantly smaller in extent than the other algorithms. Here the ARPS microphysical parameterizations do not appear to accurately predict the location of liquid water in the model domain.

The resulting icing forecasts based solely on the explicit calculations of supercooled liquid water (cloud and rain water) within the model may therefore show limited performance until the deficiencies within the microphysical processes in the model are eliminated. In addition we need to remember that the only water content adjustments incorporated in the model initialization was in the form of low resolution radar reflectivity, at a limited number of elevation angles. The inclusion of higher resolution radar data, better covering the full domain, and the inclusion of other "cloud" analysis information (i.e. satellite data) may improve the model starting point, and thus may improve the models ability to predict liquid water content. This would hopefully lead to an improvement in the ARPS SLW and LAPS algorithm performance.

As a final comparison of the icing forecasts to the pilot reports of in-flight icing, we wish to see if the algorithms have better ability to predict significantly forced icing

events (i.e. those associated with more severe intensity levels). The ability of an algorithm to correctly forecast the "worst" icing events would be of greater importance to the forecasting and flying community. From Figures 3.5 and 3.6 we saw that there were a reasonable number of icing reports throughout the forecast period of the type "light-moderate" and greater (100). Figures 4.59 to 4.60 show the average detection rates of "yes" reports when in this category range, POD_{mog}, as well as the AE and VE calculations using area and volume coverages only associated with the MOG reports. These results indicate no real improvement in the ability of the algorithms to predict the more severe events over their ability to detect more routine trace and light icing regions. In fact, there is a decrease in the Tremblay detection rate versus that when using all icing reports, meaning it actually does worse at forecasting the areas where a better forecast is most desirable. No significant trends were seen in the examination of the moderate or greater results on a day by day, or hour by hour basis. The results (not shown) were in general agreement with those based on the inclusion of all icing reports. It is believed that a more extensive data set of moderate or greater reports may show features not identifiable in the sample available for this work.

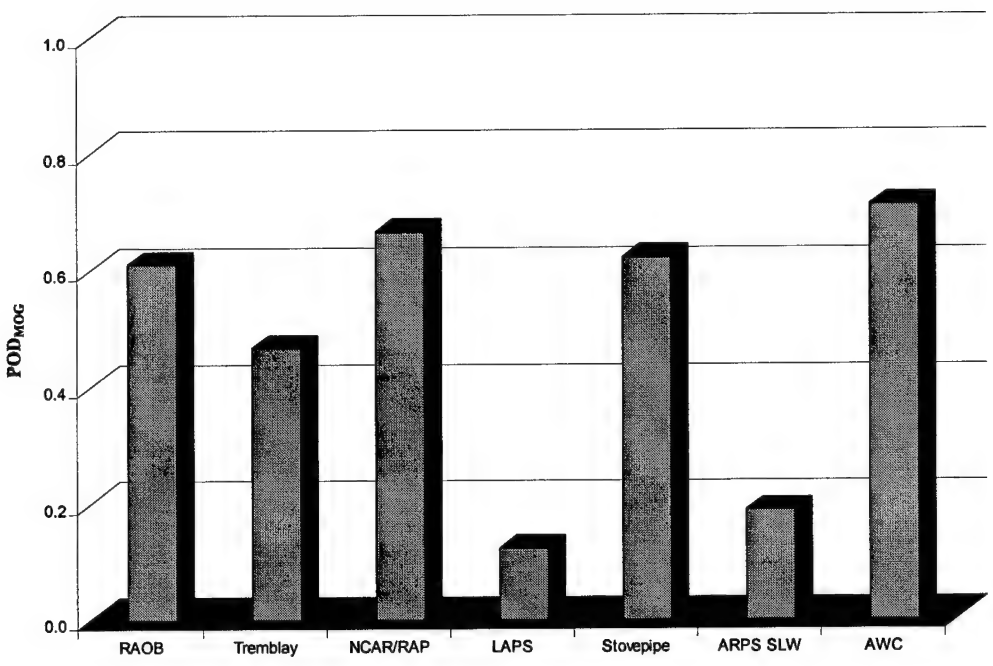


Figure 4.59: Mean percentage of detection for all moderate or greater PIREPs, POD_{mog}, during the entire forecast period.

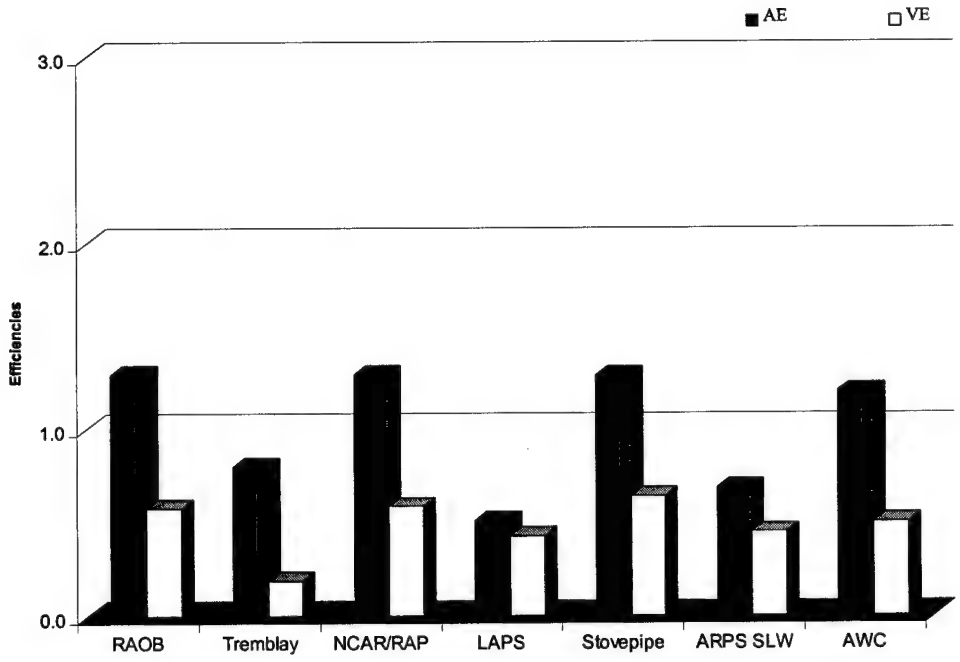


Figure 4.60: Mean Area Efficiency (AE) and Volume Efficiency (VE) for moderate or greater (MOG) PIREPs during the entire forecast period.

Chapter 5: Conclusions and Future Directions

To evaluate of the ability of a mesoscale model to provide useful forecast information to an Air Force base weather station, a three year project called COMET-Tinker was established at the University of Oklahoma. The thrust of this research effort was to incorporate the production of aviation impact variable forecasts into the regular suite of forecasting information normally generated using the ARPS forecast model. A real-time forecasting period was performed in late 1997 and early 1998 designed to provide a set of forecasts to use in the evaluation and assessment of the ARPS, and mesoscale models in general, to BWS operations.

Results obtained from the verification of the forecast output of the ARPS model against surface observations, profiler measurements, and upper level "analysis" fields were presented. In general the ARPS forecast errors were consistent with known limitations and errors already evident in the formulation of the specific version of the model used at the time of forecast generation. Temperature biases were found to exist in both positive and negative directions from ± 1 °C at model initialization, reaching ± 4-5 °C by the 09 hour forecast. Dew point errors were consistently positive (moist) throughout the set of forecasts, but generally remaining below 4 °C by the end of the forecast. These temperature and dew point biases were shown to extend well into the forecast domain. Especially for the dew points, where average biases of 3-5 °C were evident up to the 500 mb level. These errors were not identified to be of such a sufficient magnitude that use of the forecast output to create the AIV forecasts was ill-advised. In the future, a more rigorous verification procedure, including the removal of boundary

regions from consideration (Warner 1998), will provide a better "true assessment" of the actual performance of the model physics.

An intercomparison of various icing AIV forecast algorithms available to the duty forecast during the operational forecasting period was also performed. Detection rates based on pilot reports of in flight conditions and the spatial characteristics of the icing forecast by each method were examined to determine a relative order of efficiency and quality. Detection rates of observed icing conditions ranged from 20 to 60%, with the set of "synoptically" based threshold algorithms performing significantly better, even when the spatial extent of the icing regions forecast was factored in. These relationships held for comparison of all reported icing events, as well as for the subset of the more operationally significant "moderate or greater" cases, where no significant improvement was evident.

In related studies conducted by Brown et al. (1997) and Carriere et al. (1997), some of the algorithms used in this work were examined against icing forecasts based on output from other numerical weather prediction models. In those studies the relative ranking of the algorithms common to all, the RAOB, AWC, and NCAR/RAP schemes, exhibited similar characteristics. Overall detection rates of both "all" icing and MOG events were relatively uniform and reasonable high in every study. The disparity between those and the "new" algorithms included in this work point to a need to better understand the microphysical processes within the model and improve the model derivation of variables which can directly be related to icing or other AIVs. The entire suite of icing algorithms is still in need of closer examination using specific cases within

the WOP97 data set and significant weather events experienced during the remaining lifetime of the COMET-Tinker project.

The individual algorithm performances show some relative ability to accurately predict the observed conditions of in-flight icing. The production of such a large number of AIV forecast products regarding a single forecast element was done to allow the evaluation of the "best" algorithm to use with the ARPS model. The generation of a redundant set of forecasts can actually confuse forecasters by providing too much information which may sometimes be contradictory in nature. In continued forecast operations it might be desirable to combine the complete set of icing algorithms into a single "generalized" or ensemble forecast, to provide some type of consensus or "most probable" forecast based on the assessments made by all the algorithms.

The icing forecast comparison also brought to attention the need to maintain an ongoing collection of forecast and related verification information to generate as large a verification data set as possible. Continuing operations of the COMET-Tinker and Hub-CAPS projects, along with intensive forecasting periods planned by CAPS in the future should provide a wealth of additional information. The inclusion of better "observed" icing indications would be helpful, but highly unlikely with the current restrictive nature of pilot reports. Current research efforts at locations around the country that are looking into the use of satellite, radar, and other remote sensing capabilities to measure icing conditions may someday allow a more complete examination of observed flight level icing and provide a better observational data source for verification of icing forecasts made by numerical weather prediction systems.

Bibliography:

Air Weather Service, 1980: Forecaster's guide on aircraft icing, Technical Report 80/002, Scott AFB, IL, 58 pp.

Albers, S. C., J. A. McGinley, D. L. Birkenheuer, and J. R. Smart, 1996: The Local Analysis and Prediction System (LAPS): Analyses of clouds, precipitation, and temperature. *Wea. Forecasting*, **Vol 11**, 273-287

Baldwin, M., R. Treadon, and S. Contorno, 1994: Precipitation type prediction using a decision tree approach with NMC's mesoscale ETA model. *10th Conf. on Numerical Weather Prediction*, Portland, OR, Amer. Meteor. Soc., 30-31

Barnes, S. L., 1973: Mesoscale objective analysis using weighted time-series observations. NOAA Tech. Memo. ERL NSSL-62, National Severe Storms Laboratory, Norman, OK, 60 pp.

Benjamin, S. G., K. J. Brundage, and L. L. Monroe, 1994: Implementation of the Rapid Update Cycle Part I: Analysis/model description. Technical Procedures Bulletin No.416., National Oceanographic and Atmospheric Administration, National Weather Service, Silver Spring, MD, 16 pp.

_____, J.M. Brown, K.J. Brundage, B.E. Schwartz, T. G. Smirnova, and T. L. Smith, 1998: The operational RUC-2., *16th Conference on Weather Analysis and Forecasting*, Phoenix, AZ, Amer. Meteor. Soc., 249-252

Bernstien, B. C., T. A. Omeron, F MacDonough, and M. K. Politovich, 1997: The relationship between aircraft icing and synoptic-scale weather conditions. *Wea. Forecasting*, **Vol 12**, 742-762

_____, _____, _____, and _____, 1998: Surface weather features associated with freezing precipitation and severe in-flight aircraft icing. *Atmos. Res.*, **Vol 46**, 57-73-762

_____, 1996: A new technique for identifying locations where supercooled large droplets are likely to exist: The Stovepipe Algorithm. *15th Conf. on Weather Analysis and Forecasting*, Norfolk, VA, Amer. Meteor. Soc., 5-8

Blaine, C. L., 1996: *Weather Support Plan*. Oklahoma City Air Logistics Center, Air Force Material Command, Tinker AFB, OK, 24 pp.

Bratseth, A. M., 1986: Statistical interpolation by means of successive corrections. Tellus, 38A, 439-447

Brewster, K., F. Carr, N. Lin, J. Straka, and J. Krause, 1994: A local analysis system for initializing real-time convective-scale models. *10th Conf. on Numerical Weather Prediction*, Portland, OR, Amer. Meteor. Soc., 596-598.

Brock, F. V., K. C. Crawford, R. L. Elliot, G. W. Cuperus, S. J. Stadler, H. L. Johnson, and M. D. Eilts, , 1995: The Oklahoma Mesonet: A technical overview. *J. Atmos. Oceanic Technol.*, **Vol 12**, 5-19

Brown, B. G., G. Thompson, R. T. Bruintjes, R. Bullock, and T. Kane, 1997: Intercomparison of in-flight icing algorithms. Part II: WISP94 Statistical verification results. *Wea. Forecasting*, **Vol 12**, 890-914

_____, and A. H. Murphy, 1996: Verification of Aircraft icing forecasts: The use of standards measures and meteorological covariates. *13th Conf. on Probability and Statistics in the Atmospheric Sciences*, San Francisco, CA, Amer. Meteor. Soc., 251-252

Cairns M. M., R. J. Miller, S. C. Albers, D. L. Birkenheuer, B. D. Jamison, C. S. Hartsough, J. L. Mahoney, A. Marroquin, P. T. McCaslin, J. E. Ramer, and J. M. Schmidt, 1993: A preliminary evaluation of aviation-impact variables derived from numerical models. National Oceanographic and Atmospheric Administration Technical Memorandum ERL FSL-5, Boulder, CO, 165 pp.

Carpenter, R. L. Jr., K. K. Droegemeier, G. M. Bassett, K. Brewster, D. E. Jahn, J. Levit, M. Xue, W. L. Qualley, and R. Strasser, 1997: Project Hub-CAPS: Stormscale NWP for commercial aviation. *7th Conf. on Aviation , Range, and Aerospace Meteorology*, Long Beach, CA, Amer. Meteor. Soc., 474-479

Carr, F. H., J. Mewes, and K. Brewster, 1996: Quantitative verification of non-hydrostatic model forecasts of convective phenomena. *18th Conference on Severe Local Storms*, Amer. Meteor. Soc., 251-252

_____, K. R. Nixon, W. B. Beasley, C. H. Pattton, M. B. Schott, J. L. Mitchell, J.D. Bonewitz, and C. L. Bjerkaas, 1996: Proposal for COMET AWS project: Evaluation of numerical stormscale forecasts at an Air Force base weather station., University of Oklahoma, 12 pp.

_____, and J. L. Mitchell, 1997: Evaluation of numerical stormscale forecasts at an Air Force base weather station: Semi-annual outreach report – January 1997. University of Oklahoma, 3 pp.

_____, and J. L. Mitchell, 1997: Evaluation of numerical stormscale forecasts at an Air Force base weather station: Annual outreach report – July 1997. University of Oklahoma, 4 pp.

_____, and J. A. Gardner, 1998: Evaluation of numerical stormscale forecasts at an Air Force base weather station: Semi-annual outreach report – January 1998. University of Oklahoma, 3 pp.

Carriere, J. M., S. Alquier, C. Le Bot, and E. Moulin, 1997: Statistical verification of forecast icing risk indices. *Meteorological Appl.*, **Vol 4**, 115-130

Cox, R., B. L. Bauer, and T. Smith, 1998: A mesoscale model comparison. *Bull. Amer. Meteor. Soc.*, **Vol 79**, 265-283

Cornell, D., C. A. Donahue, and C. Keith, 1995: A comparison of aircraft icing forecast models. Technical Note 95/004, Air Force Combat Climatology Center, Scott AFB, IL, 33 pp.

Doswell, C.A. III, R. Davies-Jones, and D. L. Keller, 1990: On summary measures of skill in rare event forecasting based on contingency tables. *Wea. Forecasting*, **Vol 5**, 576-585

Forbes, G. S., Y. Hu, B. G. Brown, B. C. Bernstein, and M. K. Politovich, 1993: Examinations of conditions in the proximity of pilot reports of icing during STORM-FEST. *5th Conf. on Aviation Weather Systems*, Vienna, VA, Amer. Meteor. Soc., 282-286

Gandin, L. S., and A. H. Murphy, 1992: Equitable skill scores for categorical forecasts. *Mon. Wea. Rev.*, **Vol 120**, 361-370

Kelsch M., and L. Wharton, 1996: Comparing PIREPs with NAWAU turbulence and icing forecasts: Issues and results. *Wea. Forecasting*, **Vol 11**, 385-390

Kessler, E., 1969, *On the Distribution and Continuity of Water Substance in Atmospheric Circulations. Meteor. Monogr.*, No. 32, Amer. Meteor. Soc., 84 pp.

Knapp, D. I., 1992: Comparison of various icing analysis and forecasting techniques. verification report, Air Force Global Weather Center, Offutt AFB, NE, 5 pp.

McCann, D. W., 1997: Five ways to produce supercooled drizzle drops. *7th Conf. on Aviation , Range, and Aerospace Meteorology*, Long Beach, CA, Amer. Meteor. Soc., 94-99

Mewes, J. J., 1997: *Quantitative verification of non-hydrostatic model forecasts*. Master's Thesis, University of Oklahoma, School of Meteorology, 205 pp.

Murphy, A. H. and R. L. Winkler, 1987: A general framework for forecast verification. *Mon. Wea. Rev.*, **Vol 115**, 1330-1338

_____, 1991: Forecast verification: Its complexity and dimensionality. *Mon. Wea. Rev.*, **Vol 119**, 1590-1601

_____, 1993: What is a good forecast? An essay on the nature of goodness in weather forecasting. *Wea. Forecasting*, **Vol 8**, 281-293

Politovich, M. K., 1989: Aircraft icing caused by large supercooled droplets. *J. Appl. Met.*, **Vol 28**, 856-868

Research Applications Program, 1998: 1996 In-Flight icing algorithm evaluation: Stovepipe, NNICE, IIDA, RAP, satellite, and AIRMETs. National Center for Atmospheric Research, Boulder, CO, 20 pp.

Ray, P. S. (ed.), 1986: *Mesoscale Meteorology and Forecasting*. Amer. Meteor. Soc., 793 pp.

Richardson, Y. P., 1993: *Verification of NMC Short Range Models Using Wind Profiler Data*. Master's Thesis, University of Oklahoma, School of Meteorology, 102 pp.

Rogers, R. R., and M. K. Yau, 1989: *A Short Course in Cloud Physics. International Series in Natural Philosophy*, Vol. 113, 290 pp.

Sand, W. R., W. A. Cooper, M. K. Politovich, and D. L. Veal, 1984: Icing conditions encountered by a research aircraft. *J. Climate Appl. Met.*, **Vol 23**, 1427-1440

Schaefer, J. T., 1990: The Critical Success Index as an indicator of warning skill. *Wea. Forecasting*, **Vol 5**, 570-575

Schultz, P, and M. K. Politovich, 1992: Toward the improvement of aircraft-icing forecasts for the continental United States. *Wea. Forecasting*, **Vol 7**, 491-500

Schwartz, B., 1996: The quantitative use of PIREPs in developing aviation weather guidance products. *Wea. Forecasting*, **Vol 11**, 371-384

Smart, J. R., 1993: Toward improved mesoscale forecasts of aircraft icing over Eastern Colorado – Experimental evaluation and technique development. Forecast Systems Laboratory, National Oceanographic and Atmospheric Administration, 10 pp.

_____, 1996: Response of a research aircraft to icing and evaluation of severity indices. *J. Aircraft*, **Vol 33**, 291-297

Smith, T. L., and S.G. Benjamin, 1998: The Combined Use of GOES Cloud Drift, ACARS, VAD, and Profiler Winds in RUC-2. *12th Conference on Numerical Weather Prediction*, Phoenix, AZ., Amer. Meteor. Soc., 297-299

Stipanuk, G. S., 1973: Algorithms for generating a skew-t, log p diagram and computing selected meteorological quantities. Atmospheric Sciences Laboratory, U.S. Army Electronics Command White Sands Missile Range, NM, 33 pp.

Thompson, G., 1995: Pilot report text decoding software – User's guide. National Center for Atmospheric Research, Boulder, CO, 19 pp.

_____, R. T. Bruintjes, B. G. Brown, and F. Hage, 1997: Intercomparison of in-flight icing algorithms. Part I: WISP94 Real-Time icing prediction and evaluation program. *Wea. Forecasting*, **Vol 12**, 878-889

Tremblay, A., A. Glazer, W. Szyrmer, G. Isaac, and I. Zawadzki, 1995: Forecasting supercooled clouds. *Mon. Wea. Rev.*, **Vol 123**, 2098-2113

_____, S. Cober, A. Glazer, G. Isaac, and J. Mailhot, 1996: An intercomparison of mesoscale forecasts of aircraft icing using SSM/I retrievals. *Wea. Forecasting*, **Vol 11**, 66-77

Warner, T. T., R. A. Peterson, R. E. Treadon, 1997: A tutorial on lateral boundary conditions as a basic and potential serious limitation to numerical weather prediction. *Bull. Amer. Meteor. Soc.*, **Vol 78**, 2599-2617

Wilks, D. S., 1995: *Statistical Methods in the Atmospheric Sciences*. Academic Press, 467 pp.

Xue, M., K. K. Droegemeier, V. Wong, A. Shapiro, and K. Brewster, 1995: *ARPS Version 4.0 User's Guide*, Center for Analysis and Prediction of Storms, University of Oklahoma, 380 pp.

Zawadski, I., L Ostiguy, and R. Laprise, 1993: Retrieval of the microphysical properties in a CASP storm integration of a numerical kinematical model. *Atmos-Ocean*, **Vol 31**, 201-233